

OM protein - protein search, using sw model

Run on: March 24, 2004, 13:07:03 ; Search time 6.99093 Seconds
(without alignments)
2909.975 Million cell updates/sec

Title: US-09-856-681A-4
Perfect score: 376
Sequence: 1 PPPAPQRVDSIQVHSSQPSG.....PPKPSFAPLSTSMKPNDACT 72

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	376	100.0	72	3	AAY71461	Aay71461 Binding d
2	376	100.0	507	4	AAB92688	Aab92688 Human pro
3	376	100.0	562	4	AAM93444	Aam93444 Human pol
4	376	100.0	562	4	AAB94104	Aab94104 Human pro
5	376	100.0	574	4	AAB95317	Aab95317 Human pro
6	376	100.0	699	4	AAB95139	Aab95139 Human pro
7	376	100.0	863	4	ABG04066	Abg04066 Novel hum
8	376	100.0	974	2	AAW64221	Aaw64221 Human sec
9	376	100.0	975	4	AAB90731	Aab90731 Human CJ1

10	376	100.0	1030	3	AAAY71460	Aay71460	Human sem
11	376	100.0	1047	6	ADA23362	Ada23362	Human SEC
12	163.5	43.5	451	4	AAB94239	Aab94239	Human pro
13	163.5	43.5	464	4	AAB94296	Aab94296	Human pro
14	163.5	43.5	474	6	ABU11724	Abu11724	Human MDD
15	163.5	43.5	1017	5	AAG79413	Aag79413	CADHP-2,
16	163.5	43.5	1032	5	ABG79175	Abg79175	Human sem
17	163.5	43.5	1035	5	ABG79173	Abg79173	Human sem
18	163.5	43.5	1086	4	AAG63213	Aag63213	Amino aci
19	163.5	43.5	1088	5	ABG79177	Abg79177	Human sem
20	86	22.9	873	6	ABG74448	Abg74448	Rat germi
21	80.5	21.4	374	4	AAM52309	Aam52309	Murine zy
22	80.5	21.4	564	4	AAM52303	Aam52303	Murine zy
23	80.5	21.4	564	5	ABB57260	Abb57260	Mouse isc
24	79.5	21.1	1638	7	ADC31236	Adc31236	Human nov
25	77.5	20.6	693	5	ABP69529	Abp69529	Human pol
26	77.5	20.6	1111	5	ABB93287	Abb93287	Herbicida
27	75.5	20.1	497	7	ADD14148	Add14148	Human src
28	75.5	20.1	735	4	ABB66233	Abb66233	Drosophil
29	75.5	20.1	745	4	ABB68321	Abb68321	Drosophil
30	75.5	20.1	4873	6	ABO14747	Abol4747	Novel hum
31	75	19.9	95	4	AAU46062	Aau46062	Propionib
32	75	19.9	95	6	ABM42581	Abm42581	Propionib
33	74.5	19.8	449	4	ABB60133	Abb60133	Drosophil
34	74.5	19.8	1389	6	ABU42227	Abu42227	Protein e
35	74.5	19.8	1475	4	ABB71451	Abb71451	Drosophil
36	74	19.7	621	7	ADE55440	Ade55440	Rat Prote
37	74	19.7	621	7	ADD45187	Add45187	Rat Prote
38	74	19.7	628	6	ABB80175	Abb80175	A. fumiga
39	74	19.7	873	4	AAM78346	Aam78346	Human pro
40	74	19.7	884	6	AAO30334	Aao30334	Human MAP
41	74	19.7	894	2	AAAY55935	Aay55935	Human KHS
42	74	19.7	894	4	AAM78345	Aam78345	Human pro
43	74	19.7	894	6	AAO30333	Aao30333	Human MAP
44	74	19.7	930	4	AAM79329	Aam79329	Human pro
45	74	19.7	930	4	AAM79330	Aam79330	Human pro

ALIGNMENTS

RESULT 1

AAAY71461

ID AAAY71461 standard; peptide; 72 AA.

XX

AC AAAY71461;

XX

DT 04-OCT-2000 (first entry)

XX

DE Binding domain of human semaphorin 6A-1.

XX

KW Human; semaphorin 6A-1; (HSA)SEMA6A-1; neuronal development; apoptosis;

KW neuronal regeneration; Ena/VASP protein family; immunomodulatory;

KW gene therapy; diagnostic agent; therapeutic agent; differentiation;

KW cytoskeletal stabilisation; plasticity.

XX

OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Binding-site 51. .56
 FT /note= "Specific binding motif for members of Ena/VASP
 FT protein family, especially Evl"
 XX
 PN WO200031252-A1.
 XX
 PD 02-JUN-2000.
 XX
 PF 26-NOV-1999; 99WO-EP009215.
 XX
 PR 26-NOV-1998; 98EP-00122441.
 XX
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 XX
 PI Behl C, Klostermann A;
 XX
 DR WPI; 2000-400065/34.
 DR N-PSDB; AAD01234.
 XX
 PT Nucleic acid coding for human semaphorin 6A-1 used as diagnostic agent,
 PT therapeutic agent, for modulating immune system, in gene therapy or for
 PT effecting differentiation, cytoskeletal stabilization and/or plasticity.
 XX
 PS Disclosure; Page 22; 53pp; English.
 XX
 CC The present sequence is a binding domain of transmembranous human
 CC semaphorin 6A-1 ((HSA)SEMA6A-1) which is involved in neuronal development
 CC and regeneration mechanisms during apoptosis. The binding domain shows
 CC homology to Zyxin protein and selectively binds to members of Ena/VASP
 CC protein family, especially Evl. (HSA)SEMA6A-1 is a member of protein
 CC family displaying secreted or transmembrane-based repulsive guidance cues
 CC critically involved in neuronal development. Expression of (HSA)SEMA6A-1
 CC is highest in embryonic brain and kidney and moderate in lung. The
 CC present sequence is useful as diagnostic and therapeutic agents, for
 CC modulating the immune system, in gene therapy, for effecting
 CC differentiation, cytoskeletal stabilisation and plasticity
 XX
 SQ Sequence 72 AA;

Query Match 100.0%; Score 376; DB 3; Length 72;
 Best Local Similarity 100.0%; Pred. No. 1.2e-34;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 PPPAPQVRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
 QY 61 LSTSMKPNDACT 72
 |||||||||
 Db 61 LSTSMKPNDACT 72

RESULT 2
 AAB92688
 ID AAB92688 standard; protein; 507 AA.

XX
 AC AAB92688;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:11073.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 11073; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent

CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 507 AA;

Query Match 100.0%; Score 376; DB 4; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.3e-33;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
|||||
Db 436 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 495

Qy 61 LSTSMKPNDACT 72
|||||
Db 496 LSTSMKPNDACT 507

RESULT 3
AAM93444

ID AAM93444 standard; protein; 562 AA.
XX
AC AAM93444;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 3088.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR N-PSDB; AAK94365.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 3088; 1380pp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.

XX

PS Claim 8; SEQ ID NO 14328; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX

50 Sequence 562 AA;

Query Match 100.0%; Score 376; DB 4; Length 562;
Best Local Similarity 100.0%; Pred. No. 1.5e-33;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
|||||
Db 491 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 550

```

Qy          61 LSTSMKPNDACT 72
             |||||
Db          551 LSTSMKPNDACT 562

```

RESULT 5

ID AAB95317 standard; protein; 574 AA.

XX

AC AAB95317;

DT

DE Human protein sequence SEQ ID NO:17568.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 17568; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 574 AA;

Query Match 100.0%; Score 376; DB 4; Length 574;
 Best Local Similarity 100.0%; Pred. No. 1.5e-33;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
 |||
 Db 503 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 562

Qy 61 LSTSMKPNDACT 72
 |||
 Db 563 LSTSMKPNDACT 574

RESULT 6

AAB95139

ID AAB95139 standard; protein; 699 AA.

XX

AC AAB95139;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human protein sequence SEQ ID NO:17154.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-00116126.

XX

PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX

PS Claim 8; SEQ ID NO 17154; 2537pp + Sequence Listing; English.

XX

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX

SQ Sequence 699 AA;

Query Match 100.0%; Score 376; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.9e-33;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
|
Db 628 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 687

Qy 61 LSTSMKPNDACT 72
|
Db 688 LSTSMKPNDACT 699

RESULT 7

ABG04066

ID ABG04066 standard; protein; 863 AA.

XX

AC ABG04066;

XX

DT 13-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #4057.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS68253.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 34425; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 863 AA;

Query Match 100.0%; Score 376; DB 4; Length 863;
 Best Local Similarity 100.0%; Pred. No. 2.5e-33;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 792 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 851
 Qy 61 LSTSMKPNDACT 72
 ||||||||||||
 Db 852 LSTSMKPNDACT 863

RESULT 8
 AAW64221
 ID AAW64221 standard; protein; 974 AA.
 XX
 AC AAW64221;
 XX
 DT 06-OCT-1998 (first entry)

XX
DE Human secreted protein from clone CJ145_1.
XX
KW Secreted protein; human fetal brain; nutrition; cytokine; stimulant;
KW cell proliferation; differentiation; immune system; suppressor; ligand;
KW regulator; hematopoiesis; tissue growth; activin; inhibin; haemostatic;
KW chemotaxis; chemokinetic; thrombosis; receptor; cadherin; tumour;
KW anti-inflammatory.
XX
OS Homo sapiens.
XX
PN WO9827205-A2.
XX
PD 25-JUN-1998.
XX
PF 17-DEC-1997; 97WO-US023330.
XX
PR 18-DEC-1996; 96US-00769192.
PR 13-JAN-1997; 97US-00783401.
PR 16-DEC-1997; 97US-00991872.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
PI Spaulding V, Agostino MJ;
XX
DR WPI; 1998-362774/31.
DR N-PSDB; AAV44295.
XX
PT New polynucleotides and secreted proteins - obtained from human foetal
PT brain, human adult testes, human adult brain and human adult salivary
PT gland cDNA libraries.
XX
PS Claim 17j; Page 71-74; 110pp; English.
XX
CC This sequence represents a novel secreted protein from clone CJ145_1
CC isolated from a human fetal brain cDNA library. This protein has
CC applications for nutritional use, cytokine and cell
CC proliferation/differentiation activity, immune stimulating or suppressing
CC activity, hematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombotic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity and other activities
XX
SQ Sequence 974 AA;

Query Match 100.0%; Score 376; DB 2; Length 974;
Best Local Similarity 100.0%; Pred. No. 2.9e-33;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
|||||
Db 903 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 962
Qy 61 LSTSMKPNDACT 72
|||||

RESULT 9

AAB90731

ID AAB90731 standard; protein; 975 AA.

XX

AC AAB90731;

XX

DT 07-JUN-2001 (first entry)

XX

DE Human CJ145_1 protein sequence SEQ ID 161.

XX

KW Human; secreted protein; nutrient; cytokine modulator; proliferation;

KW differentiation; immune system modulator; tissue growth; chemotactic;

KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;

KW haematopoiesis.

XX

OS Homo sapiens.

XX

PN WO200119988-A1.

XX

PD 22-MAR-2001.

XX

PF 14-SEP-2000; 2000WO-US025135.

XX

PR 17-SEP-1999; 99US-00398829.

XX

PA (GEMY) GENETICS INST INC.

XX

PI Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;

XX

DR WPI; 2001-244801/25.

DR N-PSDB; AAF98469.

XX

PT Isolated nucleic acids encoding polypeptides, useful for modulating e.g.

PT cytokine and cell proliferation/differentiation activity, the immune

PT system and hematopoiesis regulating activity.

XX

PS Disclosure; Page 487-490; 557pp; English.

XX

CC Human cDNA clones represented in AAF98374 - AAF98489 encode secreted
CC proteins AAB90667 - AAB90750. The cDNA clones are isolated from various
CC tissue types, and may be used in the prevention, treatment and diagnosis
CC of diseases associated with inappropriate protein expression. The
CC polypeptides and nucleic acids may be used as nutrients or to modulate
CC cytokine and cell proliferation/differentiation activity and may also be
CC involved in modulation of the immune system. The cDNA sequences,
CC proteins, their agonists and/or antagonists exhibit haematopoiesis
CC regulating activity; tissue growth activity; activin/inhibin activity;
CC chemotactic/chemokinetic activity; haemostatic and thrombolytic activity;
CC receptor/ligand activity; anti-inflammatory activity; haematopoiesis
CC activity; cadherin/tumour suppressor activity; and/or tumour inhibition
CC activity. Included in the invention are probes represented in AAF98490 -
CC AAF98572 which are specific for the cDNA clones encoding the secreted
CC proteins

DR N-PSDB; AAD01233.
 XX
 PT Nucleic acid coding for human semaphorin 6A-1 used as diagnostic agent,
 PT therapeutic agent, for modulating immune system, in gene therapy or for
 PT effecting differentiation, cytoskeletal stabilization and/or plasticity.
 XX
 PS Example 1; Page 29-33; 53pp; English.
 XX
 CC The present sequence is a transmembranous human semaphorin 6A-1
 CC ((HSA)SEMA6A-1) which is involved in neuronal development and
 CC regeneration mechanisms during apoptosis. Semaphorin is a family of
 CC proteins displaying secreted or transmembrane-based repulsive guidance
 CC cues critically involved in neuronal development. The present sequence
 CC was isolated from human 1-ZAP Express cDNA library which was screened
 CC using a PCR fragment amplified from human neuroblastoma cell line SK-N-MC
 CC cDNA. The (HSA)SEMA6A-1 protein contains a Zyxin-like domain that
 CC selectively binds to members of Ena/VASP protein family especially Evl.
 CC Expression of (HSA)SEMA6A-1 is highest in embryonic brain and kidney and
 CC moderate in lung. The present sequence is useful as diagnostic and
 CC therapeutic agents, for modulating the immune system, in gene therapy,
 CC for effecting differentiation, cytoskeletal stabilisation and plasticity
 XX
 SQ Sequence 1030 AA;

Query Match 100.0%; Score 376; DB 3; Length 1030;
 Best Local Similarity 100.0%; Pred. No. 3.1e-33;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
 |||
 Db 959 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 1018
 Qy 61 LSTSMKPNDACT 72
 |||
 Db 1019 LSTSMKPNDACT 1030

RESULT 11

ADA23362

ID ADA23362 standard; protein; 1047 AA.

XX

AC ADA23362;

XX

DT 20-NOV-2003 (first entry)

XX

DE Human SECX polypeptide, SEC15.

XX

KW Human; secreted polypeptide; membrane-associated polypeptide; SECX; SEC1;

KW SEC2; SEC3; SEC4; SEC5; SEC6; SEC7; SEC8; SEC9; SEC10; SEC11; SEC12;

KW SEC13; SEC14; SEC15; SECX-associated disorder; lung cancer;

KW cardiovascular disease; oncology disease; immune disorder;

KW autoimmune disease; transplant rejection; allergy; AIDS; infections;

KW inflammatory disorder; arthritis; haematopoietic disorder; skin disorder;

KW atherosclerosis; restenosis; neurological disease; Alzheimer's disease;

KW trauma; wounds; spinal cord injury; skeletal disorder; cytostatic;

KW antiinflammatory; immunosuppressive; anti-HIV; antiarthritic;

KW antiarteriosclerotic; cardiant; neuroprotective; nootropic; vulnerary;

KW antiallergic; cardiant; dermatological.
 XX
 OS Homo sapiens.
 XX
 PN US2003054514-A1.
 XX
 PD 20-MAR-2003.
 XX
 PF 19-SEP-2001; 2001US-00957187.
 XX
 PR 09-MAR-1999; 99US-0123667P.
 PR 04-JAN-2000; 2000US-0174485P.
 PR 08-MAR-2000; 2000US-00520781.
 PR 19-SEP-2000; 2000US-0233798P.
 PR 20-SEP-2000; 2000US-0234082P.
 XX
 PA (SHIM/) SHIMKETS R A.
 PA (LARO/) LAROCHELLE W J.
 XX
 PI Shimkets RA, Larochelle WJ;
 XX
 DR WPI; 2003-540616/51.
 DR N-PSDB; ADA23361.
 XX
 PT New SECX nucleic acids, useful for treating or diagnosing a disorder
 PT e.g., lung cancer, cardiovascular and oncology diseases, immune disorder,
 PT and autoimmune disease.
 XX
 PS Claim 12; Page 14; 118pp; English.
 XX
 CC The present invention relates to the isolation of human secreted or
 CC membrane-associated (SECX) polypeptides designated SEC1-SEC15, and the
 CC polynucleotide sequences encoding them. Also disclosed is a method for
 CC screening for a modulator of activity or latency of SECX. The SECX
 CC polypeptide and polynucleotide sequences may be used for treating or
 CC preventing SECX-associated disorders such as lung cancer, cardiovascular
 CC and oncology diseases, immune disorders, autoimmune diseases, transplant
 CC rejection, allergy, AIDS, infections, inflammatory disorders, arthritis,
 CC haematopoietic disorders, skin disorders, atherosclerosis, restenosis,
 CC neurological diseases (e.g. Alzheimer's disease), trauma, wounds, spinal
 CC cord injuries, and skeletal disorders. The present sequence represents a
 CC SECX polypeptide of the invention.
 XX
 SQ Sequence 1047 AA;

Query Match 100.0%; Score 376; DB 6; Length 1047;
 Best Local Similarity 100.0%; Pred. No. 3.2e-33;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 976 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 1035
 Qy 61 LSTSMKPNDACT 72
 |||||||||
 Db 1036 LSTSMKPNDACT 1047

RESULT 12

AAB94239

ID AAB94239 standard; protein; 451 AA.

XX

AC AAB94239;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human protein sequence SEQ ID NO:14623.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-00116126.

XX

PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX

PS Claim 8; SEQ ID NO 14623; 2537pp + Sequence Listing; English.

XX

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by

KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
 KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
 KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
 KW psoriasis; hepatitis.
 XX
 OS Homo sapiens.
 XX
 PN WO200279449-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 27-MAR-2002; 2002WO-US009944.
 XX
 PR 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX
 DR WPI; 2003-058431/05.
 DR N-PSDB; ABX34714.
 XX
 PT New purified disease detection and treatment molecule proteins and
 PT polynucleotides, useful for diagnosing, treating or preventing cancers
 PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
 PT or hepatitis.
 XX
 PS Claim 27; SEQ ID NO 671; 339pp + Sequence Listing; English.
 XX
 CC This invention describes a novel disease detection and treatment molecule
 CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,
 CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,
 CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides
 CC and the polypeptides of the invention can be used for gene therapy,
 CC protein replacement therapy and are useful for treating a variety of
 CC diseases or conditions. These polypeptides or polynucleotides are
 CC particularly useful for diagnosing, treating or preventing cell
 CC proliferative disorders (e.g. cancers including adenocarcinoma,
 CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
 CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
 CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
 CC hepatitis. ABU11450-ABU11845 represent the MDDT polynucleotides encoded
 CC by ABU11450-ABU11845, described in the disclosure of the invention. NOTE:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 474 AA;

Query Match 43.5%; Score 163.5; DB 6; Length 474;
Best Local Similarity 50.7%; Pred. No. 9.5e-10;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
| | :|| || | :| |||: :| | :| ||||| |||||
Db 403 PTPTGAKVDYIQ-----GTPVSVHLQPSLSRQSSYTSNGTLPRTGLKRTPSLKPDVPP 455

Qy 55 KPSFAPLSTSMKP 67
| || | : | : |
Db 456 KPSFVPQTPSVRP 468

RESULT 15

AAG79413

ID AAG79413 standard; protein; 1017 AA.

XX

AC AAG79413;

XX

DT 25-OCT-2002 (first entry)

XX

DE CADHP-2, Incyte ID No: 7596315CD1.

XX

KW Human; cell adhesion protein; CADHP; AIDS; Alzheimer's disease;

KW acquired immunodeficiency syndrome; thymic dysplasia; epilepsy;

KW renal tubular acidosis; congenital glaucoma; cancer; atherosclerosis;

KW Parkinson's disease.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 1. .603

FT /label= Semaphorin_domain

FT /note= "Identified by BLAST-DOMO"

FT Peptide 1. .20

FT /label= Signal_peptide

FT /note= "Identified by HMMER"

FT Domain 4. .21

FT /label= Transmembrane_domain

FT /note= "Identified by TMAP, N-terminal domain is

FT cytoplasmic"

FT Modified-site 22

FT /note= "Potentially phosphorylated"

FT Modified-site 49

FT /note= "Potentially phosphorylated"

FT Modified-site 51

FT /note= "Potentially glycosylated"

FT Domain 59. .477

FT /label= Semaphorin_domain

FT /note= "Identified by HMMER-PFAM"

FT Binding-site 67. .182

FT /label= Semaphorin_protein_precursor_receptor

FT /note= "Identified by BLAST-PRODOM"

FT	Modified-site	70
FT		/note= "Potentially phosphorylated"
FT	Modified-site	97
FT		/note= "Potentially phosphorylated"
FT	Modified-site	151
FT		/note= "Potentially phosphorylated"
FT	Binding-site	161. .300
FT		/label= Semaphoirn_protein_precursor_receptor
FT		/note= "Identified by BLAST-PRODOM"
FT	Modified-site	187
FT		/note= "Potentially phosphorylated"
FT	Modified-site	201
FT		/note= "Potentially phosphorylated"
FT	Modified-site	210
FT		/note= "Potentially phosphorylated"
FT	Binding-site	249. .476
FT		/label= Semaphoirn_protein_precursor_receptor
FT		/note= "Identified by BLAST-PRODOM"
FT	Modified-site	266
FT		/note= "Potentially phosphorylated"
FT	Modified-site	283
FT		/note= "Potentially glycosylated"
FT	Modified-site	299
FT		/note= "Potentially phosphorylated"
FT	Modified-site	332
FT		/note= "Potentially phosphorylated"
FT	Modified-site	381
FT		/note= "Potentially phosphorylated"
FT	Modified-site	435
FT		/note= "Potentially glycosylated"
FT	Modified-site	459
FT		/note= "Potentially phosphorylated"
FT	Modified-site	461
FT		/note= "Potentially glycosylated"
FT	Modified-site	513
FT		/note= "Potentially phosphorylated"
FT	Modified-site	520
FT		/note= "Potentially phosphorylated"
FT	Modified-site	576
FT		/note= "Potentially phosphorylated"
FT	Domain	602. .630
FT		/label= Transmembrane_domain
FT		/note= "Identified by TMAP, N-terminal domain is cytoplasmic"
FT	Modified-site	650
FT		/note= "Potentially phosphorylated"
FT	Modified-site	678
FT		/note= "Potentially phosphorylated"
FT	Modified-site	687
FT		/note= "Potentially phosphorylated"
FT	Modified-site	688
FT		/note= "Potentially phosphorylated"
FT	Modified-site	734
FT		/note= "Potentially phosphorylated"
FT	Modified-site	736
FT		/note= "Potentially phosphorylated"
FT	Modified-site	745

FT /note= "Potentially phosphorylated"
 FT Modified-site 749
 FT /note= "Potentially phosphorylated"
 FT Modified-site 776
 FT /note= "Potentially glycosylated"
 FT Modified-site 782
 FT /note= "Potentially glycosylated"
 FT Modified-site 808
 FT /note= "Potentially phosphorylated"
 FT Modified-site 809
 FT /note= "Potentially phosphorylated"
 FT Modified-site 822
 FT /note= "Potentially phosphorylated"
 FT Modified-site 858
 FT /note= "Potentially phosphorylated"
 FT Modified-site 886
 FT /note= "Potentially phosphorylated"
 FT Modified-site 900
 FT /note= "Potentially phosphorylated"
 FT Modified-site 911
 FT /note= "Potentially glycosylated"
 FT Modified-site 913
 FT /note= "Potentially phosphorylated"
 FT Modified-site 978
 FT /note= "Potentially glycosylated"
 FT Binding-site 988. .1017
 FT /label= Semaphoirn_protein_precursor_receptor
 FT /note= "Identified by BLAST-PRODOM"
 FT Modified-site 991
 FT /note= "Potentially phosphorylated"
 FT Modified-site 1008
 FT /note= "Potentially phosphorylated"
 XX
 PN WO200259312-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 18-DEC-2001; 2001WO-US049206.
 XX
 PR 18-DEC-2000; 2000US-0256542P.
 PR 22-DEC-2000; 2000US-0259604P.
 PR 05-JAN-2001; 2001US-0260101P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Duggan BM, Xu Y, Lee EA, Lee S, Lu DAM, Warren BA, Yue H;
 PI Gietzen KJ, Honchell CD, Burford N, Baughn MR, Tang TY, Hillman JL;
 PI Gandhi AR, Kallick DA, Bandman O, Graul RC, Walia NK, Lu Y;
 PI Ramkumar J, Yao MG, Lal PG;
 XX
 DR WPI; 2002-590826/63.
 DR N-PSDB; ABA00055.
 XX
 PT New human cell adhesion proteins (CADHP) useful for treating, diagnosing
 PT and preventing diseases or conditions associated with the aberrant CADPH
 PT expression e.g. cancer, acquired immunodeficiency syndrome, Alzheimer's
 PT disease and epilepsy.

XX

PS Claim 1; Page 115-17; 149pp; English.

XX

The sequences given in AAG79412-21 are novel human cell adhesion proteins (CADHP). The CADHP polypeptides and polynucleotides are useful in treating, diagnosing and preventing diseases or conditions associated with the decreased expression or overexpression of CADHP, e.g. immune system (acquired immunodeficiency syndrome, thymic dysplasia), neurological (Alzheimer's disease, Parkinson's disease, epilepsy), developmental (renal tubular acidosis, congenital glaucoma) and cell proliferative (cancer, atherosclerosis) disorders. They are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of CADHP. The CADHP or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide. This protein shows homology to mouse semaphorin VIa

XX

SO Sequence 1017 AA;

Query Match 43.5%; Score 163.5; DB 5; Length 1017;

Best Local Similarity 50.7%; Pred. No. 2.4e-09;

Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

QY 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDPVPP 54

Db 946 PTPTGAKVDYIO-----GTPVSVHLOPSLSRQSSYTSNGTLPRTGLKRTPSLKPDVPP 998

Qy 55 KPSFAPLSTSMKP 67

$$\begin{array}{ccccccc} | & | & | & | & | & : & | : : | \\ \hline \end{array}$$

Db 999 KPSFVPOTPSVRP 1011

Search completed: March 24, 2004, 13:14:22

Job time : 7.99093 secs

OM protein - protein search, using sw model

Run on: March 24, 2004, 13:12:28 ; Search time 2.28675 Seconds
(without alignments)
1625.481 Million cell updates/sec

Title: US-09-856-681A-4
Perfect score: 376
Sequence: 1 PPPAPQRVDSIQVHSSQPSG.....PPKPSFAPLSTSMKPNDACT 72

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					
Result		Query					
No.	Score	Match	Length	DB	ID	Description	
1	163.5	43.5	429	4	US-09-653-274-9	Sequence 9, Appli	
2	163.5	43.5	1070	4	US-09-653-274-8	Sequence 8, Appli	
3	163.5	43.5	1086	4	US-09-653-274-4	Sequence 4, Appli	
4	74	19.7	894	4	US-09-688-188B-18	Sequence 18, Appl	
5	74	19.7	894	4	US-09-291-417D-18	Sequence 18, Appl	
6	72	19.1	480	3	US-09-189-035-5	Sequence 5, Appli	
7	72	19.1	480	3	US-09-382-086-5	Sequence 5, Appli	
8	71.5	19.0	103	4	US-09-489-039A-10263	Sequence 10263, A	
9	71.5	19.0	668	4	US-09-277-431A-2	Sequence 2, Appli	
10	71.5	19.0	1004	3	US-08-916-352-2	Sequence 2, Appli	
11	71.5	19.0	1023	4	US-10-164-595-20	Sequence 20, Appl	

12	71.5	19.0	1070	4	US-10-164-595-18	Sequence 18, Appl
13	71.5	19.0	1073	4	US-10-164-595-22	Sequence 22, Appl
14	69.5	18.5	1125	4	US-09-513-783A-152	Sequence 152, App
15	69.5	18.5	1610	4	US-09-513-783A-22	Sequence 22, Appl
16	69	18.4	400	4	US-09-252-991A-32313	Sequence 32313, A
17	68.5	18.2	625	4	US-09-196-270-6	Sequence 6, Appli
18	68	18.1	1149	3	US-08-560-005-5	Sequence 5, Appli
19	68	18.1	1149	3	US-09-418-540-5	Sequence 5, Appli
20	68	18.1	1149	4	US-09-969-528-5	Sequence 5, Appli
21	68	18.1	1253	1	US-08-252-966B-12	Sequence 12, Appl
22	68	18.1	1261	1	US-08-252-966B-18	Sequence 18, Appl
23	67	17.8	169	4	US-09-252-991A-30563	Sequence 30563, A
24	67	17.8	2035	1	US-08-046-585-5	Sequence 5, Appli
25	67	17.8	2035	1	US-08-393-703-5	Sequence 5, Appli
26	67	17.8	2035	5	PCT-US93-11721-5	Sequence 5, Appli
27	66.5	17.7	503	4	US-09-599-287A-2	Sequence 2, Appli
28	66.5	17.7	507	4	US-09-599-287A-24	Sequence 24, Appl
29	65.5	17.4	167	4	US-09-252-991A-32720	Sequence 32720, A
30	65.5	17.4	366	4	US-09-252-991A-32385	Sequence 32385, A
31	65.5	17.4	457	4	US-09-355-214-5	Sequence 5, Appli
32	65.5	17.4	1142	2	US-08-993-118-7	Sequence 7, Appli
33	65.5	17.4	1142	3	US-08-845-528C-7	Sequence 7, Appli
34	65.5	17.4	1142	3	US-09-061-709-2	Sequence 2, Appli
35	65.5	17.4	1142	4	US-09-066-281B-7	Sequence 7, Appli
36	65.5	17.4	1142	4	US-09-899-651-2	Sequence 2, Appli
37	65.5	17.4	1142	4	US-09-468-433C-7	Sequence 7, Appli
38	65.5	17.4	1142	4	US-09-392-714-26	Sequence 26, Appl
39	65.5	17.4	1297	3	US-09-540-245A-17	Sequence 17, Appl
40	65	17.3	70	4	US-09-006-428A-15	Sequence 15, Appl
41	65	17.3	143	4	US-09-252-991A-22066	Sequence 22066, A
42	65	17.3	351	3	US-08-466-465-6	Sequence 6, Appli
43	64.5	17.2	433	3	US-09-046-158A-2	Sequence 2, Appli
44	64.5	17.2	897	1	US-07-960-389-2	Sequence 2, Appli
45	64.5	17.2	1088	4	US-09-233-857-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-653-274-9

; Sequence 9, Application US/09653274

; Patent No. 6635742

; GENERAL INFORMATION:

; APPLICANT: Boyle, Bryan J

; APPLICANT: Yeung, George Y

; APPLICANT: Arterburn, Matthew C

; APPLICANT: Mize, Nancy K

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Drmanac, Radoje T

; TITLE OF INVENTION: Methods and Maaterials Relating to Semaphorin-Like

; TITLE OF INVENTION: Polypeptides and Polynucleotides

; FILE REFERENCE: HYS-23

; CURRENT APPLICATION NUMBER: US/09/653,274

; CURRENT FILING DATE: 2000-08-31

; PRIOR APPLICATION NUMBER: 09/491,404

; PRIOR FILING DATE: 2000-01-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-653-274-9

Query Match 43.5%; Score 163.5; DB 4; Length 429;
Best Local Similarity 50.7%; Pred. No. 4.8e-10;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDPVPP 54
| | :|| || | |:| ||||: :| | |:|||||||
Db 358 PTPTGAKVDYIQ-----GTPVSVHLQPSLSRQSSYTSNGTLPRTGLKRTPSLKPDPVPP 410

Qy 55 KPSFAPLSTSMKP 67
|||| | :|:|
Db 411 KPSFVPQTPSVRP 423

RESULT 2

US-09-653-274-8
; Sequence 8, Application US/09653274
; Patent No. 6635742
; GENERAL INFORMATION:
; APPLICANT: Boyle, Bryan J
; APPLICANT: Yeung, George Y
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Mize, Nancy K
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Methods and Maaterials Relating to Semaphorin-Like
; TITLE OF INVENTION: Polypeptides and Polynucleotides
; FILE REFERENCE: HYS-23
; CURRENT APPLICATION NUMBER: US/09/653,274
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1070
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-653-274-8

Query Match 43.5%; Score 163.5; DB 4; Length 1070;
Best Local Similarity 50.7%; Pred. No. 1.5e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDPVPP 54
| | :|| || | |:| ||||: :| | |:|||||||
Db 999 PTPTGAKVDYIQ-----GTPVSVHLQPSLSRQSSYTSNGTLPRTGLKRTPSLKPDPVPP 1051

Qy 55 KPSFAPLSTSMKP 67
|||| | : |::|
Db 1052 KPSFVPQTPSVRP 1064

RESULT 3

US-09-653-274-4
; Sequence 4, Application US/09653274
; Patent No. 6635742
; GENERAL INFORMATION:
; APPLICANT: Boyle, Bryan J
; APPLICANT: Yeung, George Y
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Mize, Nancy K
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Methods and Maaterials Relating to Semaphorin-Like
; TITLE OF INVENTION: Polypeptides and Polynucleotides
; FILE REFERENCE: HYS-23
; CURRENT APPLICATION NUMBER: US/09/653,274
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-653-274-4

Query Match 43.5%; Score 163.5; DB 4; Length 1086;
Best Local Similarity 50.7%; Pred. No. 1.5e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDPVPP 54
| | :|| | | | :| | | | :| | | :|||||||
Db 1015 PTPTGAKVDYIQ-----GTPVSVHLQPSLSRQSSYTSNGTLPRTGLKRTPSLKPDPVPP 1067

Qy 55 KPSFAPLSTSMKP 67
|||| | : |::|
Db 1068 KPSFVPQTPSVRP 1080

RESULT 4

US-09-688-188B-18
; Sequence 18, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/09/688,188B
; CURRENT FILING DATE: 2000-10-16

; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-188B-18

Query Match 19.7%; Score 74; DB 4; Length 894;
Best Local Similarity 32.8%; Pred. No. 7.6;
Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;

Qy 1 PPPAPQRVDSI----QVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKP 56
||| | : || :|||: | |: | | || || | || |
Db 432 PPPLPPKPKSIFIPQEMHSTEDENQG-TIKRCP-----MSGSPAKPSQVPPRPPPP 481

Qy 57 SFAP 60
|
Db 482 RLPP 485

RESULT 5

US-09-291-417D-18
; Sequence 18, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-417D-18

Query Match 19.7%; Score 74; DB 4; Length 894;
Best Local Similarity 32.8%; Pred. No. 7.6;
Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;

Qy 1 PPPAPQRVDSI----QVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKP 56
||| | : || :|||: | |: | | || || | || |
Db 432 PPPLPPKPKSIFIPQEMHSTEDENQG-TIKRCP-----MSGSPAKPSQVPPRPPPP 481

Qy 57 SFAP 60
|

Db 482 RLPP 485

RESULT 6

US-09-189-035-5
; Sequence 5, Application US/09189035
; Patent No. 6020165
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS
; FILE REFERENCE: PF-0638 US
; CURRENT APPLICATION NUMBER: US/09/189,035
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g2245671
US-09-189-035-5

Query Match 19.1%; Score 72; DB 3; Length 480;
Best Local Similarity 31.2%; Pred. No. 5.8;
Matches 24; Conservative 13; Mismatches 26; Indels 14; Gaps 4;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRS----GLKRTPSL-----KP 50
| |:| : | | | |:| | : : | | | | : | | | : : |
Db 264 PTPSPPTIG--PAPGSAPGSQYGTMTTRQ--ISRHNSTTSSTSSGGYRRTPSVTAQFSAQP 319

Qy 51 DVPPKPSFAPLSTSMKP 67
| | : : | | : |
Db 320 HVNGGPLYSQNSISIAIP 336

RESULT 7

US-09-382-086-5
; Sequence 5, Application US/09382086
; Patent No. 6201106
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS
; FILE REFERENCE: PF-0638 US
; CURRENT APPLICATION NUMBER: US/09/382,086
; CURRENT FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 09/189,035
; EARLIER FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 5

```
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g2245671
US-09-382-086-5
```

RESULT 8

```

Query Match          19.0%;  Score 71.5;  DB 4;  Length 103;
Best Local Similarity 30.8%;  Pred. No. 0.95;
Matches    24;  Conservative    7;  Mismatches    28;  Indels    19;  Gaps    3;

Qy          1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNA-----YNSLTRSGLKR-----T 45
            | | |  | : |      | | |      | | | | |      : : : | : |      |
Db          9 PAPAASRIRSPAAASSAPPAS----SRCPSSNAAPAPLPDETSATSAAGYRRRTETVQAT 64

Qy          46 PSLKPDVPPKPSFAPLST 63
            |      | |  | | :  |  |
Db          65 PGPAPDPTPSPALRPPGT 82

```

RESULT 9

```

; Patent No. 6656705
; GENERAL INFORMATION:
;   APPLICANT:  Baden, Howard P.
;   APPLICANT:  Olson, Pamela
;   APPLICANT:  Champlaud, Marie-France
;   TITLE OF INVENTION:  SCIELLIN AND USES THEREOF
;   NUMBER OF SEQUENCES:  26
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Fish & Richardson P.C.
;     STREET:    225 Franklin Street
;     CITY:      Boston
;     STATE:     MA
;     COUNTRY:   USA
;     ZIP:       02110-2804
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Diskette
;     COMPUTER:     IBM Compatible
;     OPERATING SYSTEM:  DOS
;     SOFTWARE:     FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/277,431A
;     FILING DATE:       26-MAR-1999
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  60/079,498
;     FILING DATE:       26-MAR-1998
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Myers, Louis P.
;     REGISTRATION NUMBER:  35,965
;     REFERENCE/DOCKET NUMBER:  10284/009001
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  617/542-5070
;     TELEFAX:    617/542-8906
;     TELEX:      200154
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  668 amino acids
;       TYPE:    amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     FRAGMENT TYPE:  internal
US-09-277-431A-2

```

```

Query Match          19.0%;  Score 71.5;  DB 4;  Length 668;
Best Local Similarity 30.0%;  Pred. No. 9.9;
Matches  24;  Conservative  11;  Mismatches  20;  Indels  25;  Gaps  5;

```

```

Qy      10 SIQVHSSQPSGQ-----AVTVSRQ----PSLNAYNSLTRSGL-KRTPSL 48
          |::|  || |           |  ||  |  ||: : :| :| | :
Db      127 SLEVTKLQPGGSLNANTSNTIASTSATTPVKKKRQSWFPPPPPGYNASSSTGTRRREPGV 186

Qy      49 KPDVPPKPSFAPLSTSMKPN 68
          | :||||| :|:|:  ||
Db      187 HPPIPPKPS-SPVSS---PN 202

```

```

RESULT 10
US-08-916-352-2

```

```

; Sequence 2, Application US/08916352
; Patent No. 6166191
; GENERAL INFORMATION:
; APPLICANT:  CHIRON CORPORATION
; TITLE OF INVENTION:  HUMAN POLYHOMEOTIC 1 (hph1) ACTS AS A
; TITLE OF INVENTION:  TUMOR SUPPRESSOR
; NUMBER OF SEQUENCES:  2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  CHIRON CORPORATION
; STREET:  4560 HORTON STREET
; CITY:  EMERYVILLE
; STATE:  CA
; COUNTRY:  USA
; ZIP:  94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/916,352
; FILING DATE:
; CLASSIFICATION:  435
; ATTORNEY/AGENT INFORMATION:
; NAME:  POTTER, JANE
; REGISTRATION NUMBER:  33,332
; REFERENCE/DOCKET NUMBER:  1355.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  510-923-2707
; TELEFAX:  510-655-3542
; INFORMATION FOR SEQ ID NO:  2:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  1004 amino acids
; TYPE:  amino acid
; STRANDEDNESS:  single
; TOPOLOGY:  linear
; MOLECULE TYPE:  protein
US-08-916-352-2

```

```

Query Match          19.0%; Score 71.5; DB 3; Length 1004;
Best Local Similarity 35.8%; Pred. No. 17;
Matches 24; Conservative 5; Mismatches 23; Indels 15; Gaps 3;

```

```

Qy      2 PPAPQVRVDSIQVHSSQPSGQAVTVSRQP SLNAYNSLTRSGLKRTP-SLKPDVPPKPSFAP 60
      | || : || || || |: || | : :| || || || |
Db      448 PQPPQVPPTQQVPPSQSQQAQTLVVQPMLQS-----SPLSLPPDAAPKP---P 493

Qy      61 LSTSMKP 67
      : ||
Db      494 IPIQSKP 500

```

```

RESULT 11
US-10-164-595-20
; Sequence 20, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:

```

```
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: 1U 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-595-20
```

```
Query Match          19.0%;  Score 71.5;  DB 4;  Length 1023;
Best Local Similarity 27.5%;  Pred. No. 17;
Matches 22;  Conservative 12;  Mismatches 33;  Indels 13;  Gaps 2;
```

```
Qy      4 APQRVDSIQVHSSQPSGQAV-----TVSRQPSLNAYNSLTRSGLKRTPSLKP- 50
      :| | |: ::| ||| ||| :: ||| :| :|
Db      261 SPGRPQSLLDASTSDSQAVNMNTEQSQNSIVSRIKVFEGQTNIETSGLPKKPEITPR 320

Qy      51 DVPPKPSFAPLSTSMKPND 70
      :|||: : |: | |
Db      321 SLPPKPTVSSGKPSVAPKPA 340
```

RESULT 12

```
US-10-164-595-18
; Sequence 18, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: 1U 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 1070
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-595-18
```

```
Query Match          19.0%;  Score 71.5;  DB 4;  Length 1070;
Best Local Similarity 27.5%;  Pred. No. 18;
Matches 22;  Conservative 12;  Mismatches 33;  Indels 13;  Gaps 2;
```

```
Qy      4 APQRVDSIQVHSSQPSGQAV-----TVSRQPSLNAYNSLTRSGLKRTPSLKP- 50
      :| | |: ::| ||| ||| :: ||| :| :|
Db      261 SPGRPQSLLDASTSDSQAVNMNTEQSQNSIVSRIKVFEGQTNIETSGLPKKPEITPR 320

Qy      51 DVPPKPSFAPLSTSMKPND 70
      :|||: : |: | |
Db      321 SLPPKPTVSSGKPSVAPKPA 340
```

RESULT 13

US-10-164-595-22

; Sequence 22, Application US/10164595
 ; Patent No. 6657054
 ; GENERAL INFORMATION:
 ; APPLICANT: OriGene Technologies, Inc
 ; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
 ; FILE REFERENCE: 1U 103 R1
 ; CURRENT APPLICATION NUMBER: US/10/164,595
 ; CURRENT FILING DATE: 2002-06-10
 ; NUMBER OF SEQ ID NOS: 80
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 22
 ; LENGTH: 1073
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-164-595-22

Query Match 19.0%; Score 71.5; DB 4; Length 1073;
 Best Local Similarity 27.5%; Pred. No. 18;
 Matches 22; Conservative 12; Mismatches 33; Indels 13; Gaps 2;

Qy 4 APQRVDSIQVHSSQPSGQAV-----TVSRQPSLNAYNSLTRSGLKRTPSLKP- 50
 :| | |: ::| ||| ||| :: ||| :| :|
 Db 264 SPGRPQSLLDNASTSDSQAVNMIMNTEQSQNSIVSRIKVFEGQTNietsGLPKKPEITPR 323
 Qy 51 DVPPKPSFAPLSTSMKPND 70
 :||||: : |: | |
 Db 324 SLPPKPTVSSGKPSVAPKPA 343

RESULT 14

US-09-513-783A-152

; Sequence 152, Application US/09513783A
 ; Patent No. 6416959
 ; GENERAL INFORMATION:
 ; APPLICANT: Giuliano, Kenneth A.
 ; APPLICANT: Kapur, Ravi
 ; TITLE OF INVENTION: A System for Cell Based Screening
 ; FILE REFERENCE: 97-022-L1
 ; CURRENT APPLICATION NUMBER: US/09/513,783A
 ; CURRENT FILING DATE: 2000-02-25
 ; NUMBER OF SEQ ID NOS: 180
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 152
 ; LENGTH: 1125
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-513-783A-152

Query Match 18.5%; Score 69.5; DB 4; Length 1125;
 Best Local Similarity 28.6%; Pred. No. 32;
 Matches 18; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

Qy 3 PAP-QRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAPL 61
 |:| : :: : ||| ||| | :| : : |: :| |: ||
 Db 616 PSPLNLEQKETPGSQPSEPCSGVSRQEEAKAAVGVTGNDITTPPNKEPPPSPEKKAKPL 675

Qy 62 STS 64
:|:
Db 676 ATT 678

RESULT 15

US-09-513-783A-22
; Sequence 22, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1610
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: EYFP-DEVD-MAP4-EBFP construct
US-09-513-783A-22

Query Match 18.5%; Score 69.5; DB 4; Length 1610;
Best Local Similarity 28.6%; Pred. No. 49;
Matches 18; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

Qy 3 PAP-QRVD SIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAPL 61
|:| : :: : |||| |||| | :| : : |: :| |: ||
Db 862 PSPLENLEQKETPGSQPSEPCSGVSRQEEAKAAVGVTGNDITTPPNKEPPPSPEKKAKPL 921

Qy 62 STS 64
:|:
Db 922 ATT 924

Search completed: March 24, 2004, 13:17:59
Job time : 2.28675 secs

OM protein - protein search, using sw model

Run on: March 24, 2004, 13:11:23 ; Search time 2.15608 Seconds
 (without alignments)
 3212.214 Million cell updates/sec

Title: US-09-856-681A-4
 Perfect score: 376
 Sequence: 1 PPPAPQRVDSIQVHSSQPSG.....PPKPSFAPLSTSMKPNDACT 72

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : PIR_78:*
 1: pir1:*
 2: pir2:*
 3: pir3:*
 4: pir4:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	%		Query Match Length DB	ID	Description
	Score				
1	87	23.1	961 2	A55380	faciogenital dyspl
2	80.5	21.4	1322 2	A59288	myosin heavy chain
3	77.5	20.6	1111 2	T05646	hypothetical prote
4	77	20.5	175 2	T47463	serine/proline-ric
5	76.5	20.3	744 2	E86255	hypothetical prote
6	75.5	20.1	393 2	T33103	lin-1 protein - Ca
7	75.5	20.1	494 2	A42170	zinc finger protei
8	75.5	20.1	497 2	JC5076	myc-associated zin
9	74	19.7	452 2	S22199	imidazoleglycerol-
10	73.5	19.5	2282 2	T42717	DNA-binding protei
11	73	19.4	867 2	T41308	hypothetical zinc-
12	72.5	19.3	628 2	S01955	hypothetical prote
13	72.5	19.3	657 2	B84869	probable SF16 prot

14	72.5	19.3	4957	2	T03455	ALR protein - huma
15	72.5	19.3	5262	2	T03454	ALR protein - huma
16	72	19.1	459	2	A41977	retinoic acid rece
17	71.5	19.0	446	2	A42029	transcription fact
18	71.5	19.0	1522	2	T39371	transcription regu
19	71.5	19.0	2957	2	T33152	hypothetical prote
20	71	18.9	621	2	JC7278	adaptor protein co
21	71	18.9	1150	2	S58775	mypl protein - smu
22	70.5	18.8	468	2	T48615	hypothetical prote
23	70.5	18.8	1420	2	T37781	probable cytoskele
24	70	18.6	719	2	S62466	probable ATP-depen
25	70	18.6	747	2	S35546	ATP-dependent RNA
26	70	18.6	792	2	T26050	hypothetical prote
27	70	18.6	1012	2	I53172	RAE-28 - mouse
28	70	18.6	1201	2	G86441	unknown protein [i
29	69.5	18.5	331	2	B47236	zinc-finger protei
30	69.5	18.5	477	2	A47236	zinc-finger protei
31	69.5	18.5	1125	2	B41206	microtubule-associ
32	69	18.4	3942	2	T42730	Bassoon protein -
33	68.5	18.2	625	2	S48941	regulatory protein
34	68.5	18.2	1106	2	T31742	hypothetical prote
35	68.5	18.2	1188	2	S49915	extensin-like prot
36	68	18.1	428	1	TVHUEK	transforming prote
37	68	18.1	530	2	A45690	transactivator EBN
38	68	18.1	613	2	T47975	auxin response fac
39	68	18.1	736	2	T25447	hypothetical prote
40	68	18.1	963	2	T40873	probable transcrip
41	68	18.1	1172	2	T00065	hypothetical prote
42	68	18.1	1219	2	I61713	co-repressor prote
43	68	18.1	1229	2	A56068	co-repressor prote
44	68	18.1	1258	2	JC5765	inositol polyphosp
45	67.5	18.0	429	2	JC4965	elk1 protein - mou

ALIGNMENTS

RESULT 1

A55380

faciogenital dysplasia-associated protein FGD1 - human

C;Species: Homo sapiens (man)

C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 17-Mar-1999

C;Accession: A55380

R;Pasteris, N.G.; Cadle, A.; Logie, L.J.; Porteous, M.E.M.; Schwartz, C.E.;

Stevenson, R.E.; Glover, T.W.; Wilroy, R.S.; Gorski, J.L.

Cell 79, 669-678, 1994.

A;Title: Isolation and characterization of the faciogenital dysplasia (Aarskog-Scott syndrome) gene: a putative Rho/Rac guanine nucleotide exchange factor.

A;Reference number: A55380; MUID:95042764; PMID:7954831

A;Accession: A55380

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-961 <PAS>

A;Cross-references: GB:U11690; NID:g595424; PID:g595425

C;Superfamily: CDC24 homology; pleckstrin repeat homology

F;373-561/Domain: CDC24 homology <CD24>

Query Match 23.1%; Score 87; DB 2; Length 961;
Best Local Similarity 34.8%; Pred. No. 1.2;
Matches 23; Conservative 5; Mismatches 22; Indels 16; Gaps 2;

```
Qy      2 PPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 61
      | |||:      | |      | |::|      | ||| | || |||||: :
Db      127 PEGPQRL-----RSDPGPPTETPSQRP-----SPLKRAPGPKPQVPPKPSYLQM 170

Qy      62 STSMKP 67
      |
Db      171 PRMPPP 176
```

RESULT 2

A59288

myosin heavy chain Myr 8 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000

C;Accession: A59288

R;Patel, K.G.; Liu, C.; Cameron, P.L.; Cameron, R.S.

submitted to GenBank, November 1999

A;Description: Identification of a Novel Mammalian Myosin Class, XVI, in
Developing Brain.

A;Reference number: A59288

A;Accession: A59288

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1322 <PAT>

A;Cross-references: GB:AF209114; PIDN:AAF20150.1

A;Experimental source: strain Sprague-Dawley; clone KP4; cell type type 1
astrocyte

C;Superfamily: myosin motor domain homology

F;404-1132/Domain: myosin motor domain homology <MMO>

Query Match 21.4%; Score 80.5; DB 2; Length 1322;
Best Local Similarity 38.4%; Pred. No. 7.4;
Matches 28; Conservative 7; Mismatches 25; Indels 13; Gaps 4;

```
Qy      3 PAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRS-GLKRTPSLKPDVPPKPSFAP- 60
      | | |||:      : ||      || |||: : |: | || || : ||| |
Db      1248 PVPMAVDSLQALAGPS-----SRSPSLHSVFSMDDSTGL---PSPRKQPPPKPKRDPN 1298

Qy      61 --LSTSMKPNDAC 71
      || | : ||
Db      1299 TRLSASYEAVSAC 1311
```

RESULT 3

T05646

hypothetical protein F20D10.310 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999

C;Accession: T05646

R;Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.;
Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, February 1999

A;Reference number: Z15420

A;Accession: T05646
A;Molecule type: DNA
A;Residues: 1-1111 <BEV>
A;Cross-references: EMBL:AL035538
A;Experimental source: cultivar Columbia; BAC clone F20D10
C;Genetics:
A;Map position: 4
A;Introns: 139/2; 675/3
A;Note: F20D10.310

Query Match 20.6%; Score 77.5; DB 2; Length 1111;
Best Local Similarity 29.0%; Pred. No. 12;
Matches 20; Conservative 13; Mismatches 25; Indels 11; Gaps 2;

```
Qy      2 PPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPD-----VPP 54
      || : :| :: ||| | :|: | | ||:| :: : | | :||
Db      5 PPQTSK----KVRNNSGSGQTVKFARTSSGRYVSLSRDNIELSGELSGDYSNYTVHIPP 60

Qy      55 KPSFAPLST 63
      | |::|
Db      61 TPDNQPMAT 69
```

RESULT 4

T47463

serine/proline-rich protein - Arabidopsis thaliana

N;Alternate names: protein T14D3.170

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C;Accession: T47463

R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.;

Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.

submitted to the Protein Sequence Database, February 2000

A;Reference number: Z24467

A;Accession: T47463

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-175 <JOR>

A;Cross-references: EMBL:AL138649

A;Experimental source: cultivar Columbia; BAC clone T14D3

C;Genetics:

A;Map position: 3

A;Note: T14D3.170

Query Match 20.5%; Score 77; DB 2; Length 175;
Best Local Similarity 27.0%; Pred. No. 1.8;
Matches 20; Conservative 14; Mismatches 26; Indels 14; Gaps 2;

```
Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDV----PPKP 56
      | ||:| || :|:| || | ::|| | :: : :|: | |
Db      27 PAPSPDLADSPLIHASPPS-----KLGSHNSPAESPIEYSSPPEPETEHSPPSPSP 76

Qy      57 SFAPLSTSMKPND 70
      : :| : ||:
Db      77 ANSPSVSPPLPND 90
```

E86255

Query Match 20.3%; Score 76.5; DB 2; Length 744;
Best Local Similarity 30.3%; Pred. No. 9.6;
Matches 20; Conservative 14; Mismatches 23; Indels 9; Gaps 2;

```

Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
      |||      | |:| | : :| |:: ||:      | :||:: || | ::
Db    398 PPP-----IYVYSSPPPPSSKMS--PTVRAYSPPPPSSKMSPSVRAYSPPPPPYSK 448

Qy      61 LSTSMK 66
      :| |::
Db    449 MSPSVR 454

```

T33103

```
lin-1 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jun-2000
C;Accession: T33103
R;Miller, N.; Biewald, T.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of C. elegans cosmid C37F5.
A;Reference number: Z21283
A;Accession: T33103
```

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-393 <MIL>
A;Cross-references: EMBL:AF067606; PIDN:AAC17530.1; GSPDB:GN00022; CESP:C37F5.1
A;Experimental source: strain Bristol N2; clone C37F5
C;Genetics:
A;Gene: lin-1; CESP:C37F5.1
A;Map position: 4
A;Introns: 94/3; 188/2; 330/1
C;Superfamily: elk-1 transforming protein; ets DNA-binding domain homology
F;26-105/Domain: ets DNA-binding domain homology <ETS>

Query Match 20.1%; Score 75.5; DB 2; Length 393;
Best Local Similarity 30.9%; Pred. No. 6;
Matches 30; Conservative 4; Mismatches 32; Indels 31; Gaps 5;

```

Qy      1 PPPAPQR-----VDSIQVHS-SQPS-----GQAVTVSRQPSLNAYNSL 37
          ||| ||           ||  || | ||           :| | || ::
Db      151 PPPPQNPRGNTDFSALLGTDSPTHSVSTPSPTDSVCSPSSSVASSATPSTSSPVDE 210

Qy      38 TRSGLKRTPSLKPD-----VPPKPSFAPLSTSMKPN 68
          :|  ||  || |           || |  |  ||||
Db      211 SRQCRKR--SLSPSTTSSTTAPPPPPQPPTKKGMPN 245

```

RESULT 7

A42170

zinc finger protein MAZ - human (fragment)

N;Alternate names: MYC-associated zinc finger protein MAZ; zinc finger protein ZF87

C;Species: Homo sapiens (man)

C;Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 03-Jun-1996

C;Accession: A42170; A46153

R;Pyrç, J.J.; Moberg, K.H.; Hall, D.J.

Biochemistry 31, 4102-4110, 1992

A;Title: Isolation of a novel cDNA encoding a zinc-finger protein that binds to two sites within the c-myc promoter.

A;Reference number: A42170; MUID:92232709; PMID:1567856

A;Accession: A42170

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-494 <PYR>

A;Cross-references: GB:J05371

A;Note: it is uncertain whether Met-18 is the initiator or whether translation is initiated upstream to the sequenced region

R;Bossone, S.A.; Asselin, C.; Patel, A.J.; Marcu, K.B.

Proc. Natl. Acad. Sci. U.S.A. 89, 7452-7456, 1992

A;Title: MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences regulating transcriptional initiation and termination.

A;Reference number: A46153; MUID:92366479; PMID:1502157

A;Accession: A46153

A;Molecule type: mRNA

A;Residues: 18-417, 'L', 419-494 <BOS>

A;Cross-references: GB:M94046

A;Experimental source: HeLa cells

A;Note: sequence extracted from NCBI backbone (NCBIN:110666, NCBIP:110667)

C;Keywords: DNA binding; zinc finger

F;113-125/Region: alanine-rich
F;174-183/Region: alanine-rich
F;207-230/Region: zinc finger
F;296-318/Region: zinc finger
F;324-346/Region: zinc finger
F;354-368/Region: zinc finger
F;373-405/Region: zinc finger
F;409-430/Region: zinc finger
F;452-468/Region: alanine-rich

Query Match 20.1%; Score 75.5; DB 2; Length 494;
Best Local Similarity 25.8%; Pred. No. 7.7;
Matches 23; Conservative 15; Mismatches 32; Indels 19; Gaps 3;

Qy 1 PPPAPQ-----RVDSIQV-----HSSQPSGQAVTVSRQPSLNAYNSLTRSGLK 43
||| || :|| :| :: :| : |: |:: :||
Db 86 PPPTPQAPAAEPLQVDLLPVLAAAQESAAAAAAAAAAAAVAAAPPAPAAASTVDTAALK 145

Qy 44 RTPSLKPDVPPKPSFAPLSTSMKPNDACT 72
: |: | || | || : : | ||
Db 146 QPPA--PPPPPPVVSAPAAEAAPPASAAT 172

RESULT 8

JC5076

myc-associated zinc-finger protein - human

N;Alternate names: MAZ protein

C;Species: Homo sapiens (man)

C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 05-Nov-1999

C;Accession: JC5076

R;Tsutsui, H.; Sakatsume, O.; Itakura, K.; Yokoyama, K.K.

Biochem. Biophys. Res. Commun. 226, 801-809, 1996

A;Title: Members of the MAZ family: A novel cDNA clone for MAZ from human pancreatic islet cells.

A;Reference number: JC5076; MUID:96428591; PMID:8831693

A;Accession: JC5076

A;Molecule type: mRNA

A;Residues: 1-497 <TSU>

A;Cross-references: DDBJ:D85131; NID:g1752741; PIDN:BAA12728.1; PID:d1013410; PID:g1752742

A;Experimental source: pancreatic islet

C;Comment: This protein plays a role in the control of transcriptional initiation of genes for CD4 and serotonin and in termination of transcription between closely spaced human genes for complement and between the introns of the mouse gene for immunoglobulin M-D.

C;Keywords: phosphoprotein; zinc finger

F;146,204,480/Binding site: phosphate (Ser) (covalent) (by casein kinase II)

#status predicted

F;349/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 20.1%; Score 75.5; DB 2; Length 497;
Best Local Similarity 25.8%; Pred. No. 7.7;
Matches 23; Conservative 15; Mismatches 32; Indels 19; Gaps 3;

Qy 1 PPPAPQ-----RVDSIQV-----HSSQPSGQAVTVSRQPSLNAYNSLTRSGLK 43
||| || :|| :| :: :| : |: |:: :||
Db 95 PPPTPQAPAAEPLQVDLLPVLAAAQESAAAAAAAAAAAAVAAAPPAPAAASTVDTAALK 154

Qy 44 RTPSLKPDVPPKPSFAPLSTSMKPNDACT 72
 : |: | || | || : : | | |
 Db 155 QPPA--PPPPPPVVSAPAAEAAPPASAAT 181

RESULT 9

S22199
 imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19) - potato buckeye rot agent
 C;Species: Phytophthora nicotianae var. parasitica (potato buckeye rot agent)
 C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 29-Oct-1999
 C;Accession: S22199
 R;Karlovsky, P.
 submitted to the EMBL Data Library, January 1992
 A;Reference number: S22198
 A;Accession: S22199
 A;Molecule type: DNA
 A;Residues: 1-452 <KAR>
 A;Cross-references: EMBL:Z11591; NID:g3197; PIDN:CAA77675.1; PID:g3198
 C;Superfamily: imidazoleglycerol-phosphate dehydratase homology
 C;Keywords: carbon-oxygen lyase; hydro-lyase
 F;286-451/Domain: imidazoleglycerol-phosphate dehydratase homology <IPD>

Query Match 19.7%; Score 74; DB 2; Length 452;
 Best Local Similarity 34.4%; Pred. No. 9.7;
 Matches 21; Conservative 8; Mismatches 22; Indels 10; Gaps 3;

Qy 12 QVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKP-----DVPPKPSFAPLSTSM 65
 :::| || || | | : || |:: | | | |||| | : ::
 Db 111 ELHRRQPKGMAVVTGR-PRKDCAKFLTTHGIE---DLFPVQIWLEDCPPKPSPEPILLAL 166
 Qy 66 K 66
 |
 Db 167 K 167

RESULT 10

T42717
 DNA-binding protein Rc - mouse
 N;Alternate names: Ig kappa chain gene enhancer Recognition component
 C;Species: Mus musculus (house mouse)
 C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
 C;Accession: T42717
 R;Wu, L.C.; Liu, Y.; Strandtmann, J.; Mak, C.H.; Lee, B.; Li, Z.; Yu, C.Y.
 Genomics 35, 415-424, 1996
 A;Title: The mouse DNA binding protein Rc for the kappa B motif of transcription
 and for the V(D)J recombination signal sequences contains composite DNA-protein
 interaction domains and belongs to a new family of large transcriptional
 proteins.
 A;Reference number: Z22238; MUID:97001141; PMID:8812474
 A;Accession: T42717
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-2282 <WUL>
 A;Cross-references: EMBL:L46815; NID:g1377885; PID:g1377886; PIDN:AAB40884.1
 A;Experimental source: strain BALB/c; clone T1; thymocyte, brain
 C;Genetics:

A;Gene: Rc
C;Function:
A;Description: binds V(D)J recombination signal sequence and kappa B motif
C;Superfamily: HIV-EP2 enhancer-binding protein
C;Keywords: DNA recombination; transcription factor

Query Match 19.5%; Score 73.5; DB 2; Length 2282;
Best Local Similarity 38.1%; Pred. No. 63;
Matches 24; Conservative 6; Mismatches 20; Indels 13; Gaps 4;

Qy 14 HSSQPSGQAVTVSRQPSLNAYNSLTRSGL-----KRTPSLKPDVPP--KPSFAPLS-TS 64
| : | :| : || | : | || | : | :|| || ||| ||
Db 1489 HGTAPGSEALKEYAQPSSKAH----RRGLPPMSVKKEDPKEQTDLPPLAPPSSLPLSDTS 1544

Qy 65 MKP 67
||
Db 1545 PKP 1547

RESULT 11

T41308

hypothetical zinc-finger protein - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000

C;Accession: T41308

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Wedler, H.; Wambutt, R.; Wedler, E.
submitted to the EMBL Data Library, March 1998

A;Reference number: Z21986

A;Accession: T41308

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-867 <WOO>

A;Cross-references: EMBL:AL022245; PIDN:CAA18305.1; GSPDB:GN00068;

SPDB:SPCC320.03

A;Experimental source: strain 972h-; cosmid c320

C;Genetics:

A;Gene: SPDB:SPCC320.03

A;Map position: 3

C;Superfamily: GAL4 zinc binuclear cluster homology

F;71-113/Domain: GAL4 zinc binuclear cluster homology <GL4>

Query Match 19.4%; Score 73; DB 2; Length 867;
Best Local Similarity 29.4%; Pred. No. 25;
Matches 20; Conservative 11; Mismatches 35; Indels 2; Gaps 1;

Qy 2 PPAPQRVDSI--QVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFA 59
| || :: :| | | : :||:| : ||:| | | :||
Db 329 PTVNDRVSNVLPISITSEFSSVTTVPNSPATLNSYTTTSVPSGMSRHPMLMNPSTPEPSLG 388

Qy 60 PLSTSMKP 67
| |::|
Db 389 VNPSLRP 396

RESULT 12

S01955

hypothetical protein, 69K - turnip yellow mosaic virus

C;Species: turnip yellow mosaic virus, TYMV
 C;Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-2000
 C;Accession: S01955
 R;Morch, M.D.; Boyer, J.C.; Haenni, A.L.
 Nucleic Acids Res. 16, 6157-6173, 1988
 A;Title: Overlapping open reading frames revealed by complete nucleotide sequencing of turnip yellow mosaic virus genomic RNA.
 A;Reference number: S01955; MUID:88289359; PMID:3399388
 A;Accession: S01955
 A;Status: preliminary
 A;Molecule type: genomic RNA
 A;Residues: 1-628 <MOR>
 A;Cross-references: EMBL:X07441; NID:g62222; PIDN:CAA30321.1; PID:g62223
 A;Note: the authors translated the codon ACG for residue 459 as U
 C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 19.3%; Score 72.5; DB 2; Length 628;
 Best Local Similarity 29.6%; Pred. No. 19;
 Matches 21; Conservative 7; Mismatches 14; Indels 29; Gaps 3;

Qy 2 PPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTP----SLKPDV-PPKP 56
 ||||| |: :| ::|| | | : :|| | |
 Db 119 PPAPQRQHSLPLHITRPS-----RFPHHFHARRPDVLPSPV 154
 Qy 57 SFAPLSTSMKP 67
 |: | ||
 Db 155 DHGPVLTETKP 165

RESULT 13

B84869

probable SF16 protein (*Helianthus annuus*) [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002

C;Accession: B84869

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: B84869

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-657 <STO>

A;Cross-references: GB:AE002093; NID:g2281102; PIDN:AAB64038.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g43680

A;Map position: 2

C;Superfamily: *Arabidopsis thaliana* hypothetical protein T16L24.240

Query Match 19.3%; Score 72.5; DB 2; Length 657;

Best Local Similarity 35.1%; Pred. No. 20;
Matches 27; Conservative 9; Mismatches 28; Indels 13; Gaps 4;

```
Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPV-----PKP 56
      ||| |   | :| | :|: |   | || |   : |:   : || : :||   |||
Db      80 PFFRPA---SPRVASPRPTSPRVASPRVPSRA--EVPRTLSPKPPSPRAEVPRSLSPKP 134

Qy      57 ----SFAPLSTSMKPND 69
      : | | | | |
Db     135 PSRADLPRSLSPKPF 151
```

RESULT 14

T03455

ALR protein - human

C;Species: Homo sapiens (man)

C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 27-Oct-2003

C;Accession: T03455

R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano, T.; Alder, H.; Croce, C.M.; Huebner, K.; Mazo, A.; Canaani, E.

Oncogene 15, 549-560, 1997

A;Title: Structure and expression pattern of human ALR, a novel gene with strong homology to ALL-1 involved in acute leukemia, and to Drosophila trithorax.

A;Reference number: Z14954; MUID:97388474; PMID:9247308

A;Accession: T03455

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-4957 <PRA>

A;Cross-references: EMBL:AF010404; NID:g2358286; PIDN:AAC51735.1; PID:g2358287

C;Genetics:

A;Gene: ALR

A;Map position: 12

C;Superfamily: acute lymphoblastic leukemia protein, ALR type

C;Keywords: alternative splicing

Query Match 19.3%; Score 72.5; DB 2; Length 4957;
Best Local Similarity 34.7%; Pred. No. 1.8e+02;
Matches 25; Conservative 6; Mismatches 24; Indels 17; Gaps 4;

```
Qy      12 QVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRT-----PSLKPDPV-----PKP---- 56
      ::|: |||| | | |: | | :| | ||| || | |
Db     1925 ELHAKVPSGQPPNFVRSPGTGAFVG-TPSPMRFTFPQAVGEPSLKPPVPQGLPPPHGIN 1983

Qy      57 -SFAPLSTSMKP 67
      | | | |
Db     1984 SHFGPGPTLGKP 1995
```

RESULT 15

T03454

ALR protein - human

C;Species: Homo sapiens (man)

C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 27-Oct-2003

C;Accession: T03454

R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano, T.; Alder, H.; Croce, C.M.; Huebner, K.; Mazo, A.; Canaani, E.

Oncogene 15, 549-560, 1997

A;Title: Structure and expression pattern of human ALR, a novel gene with strong homology to ALL-1 involved in acute leukemia, and to Drosophila trithorax.
A;Reference number: Z14954; MUID:97388474; PMID:9247308
A;Accession: T03454
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-5262 <PRA>
A;Cross-references: EMBL:AF010403; NID:g2358284; PIDN:AAC51734.1; PID:g2358285
C;Genetics:
A;Gene: ALR
A;Map position: 12
C;Superfamily: acute lymphoblastic leukemia protein, ALR type
C;Keywords: alternative splicing

Query Match 19.3%; Score 72.5; DB 2; Length 5262;
Best Local Similarity 34.7%; Pred. No. 2e+02;
Matches 25; Conservative 6; Mismatches 24; Indels 17; Gaps 4;

```

Qy      12 QVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRT-----PSLKPDVP----PKP---- 56
          ::|: ||| | | | | : | | : | | ||| | | | |
Db      2230 ELHAKVPSGQPPNFVRSPGTGAFVG-TPSPMRFTFPQAVGEP SLKPPVPQ PGLPPPHGIN 2288

Qy      57 -SFAPLSTSMKP 67
          | | | |
Db      2289 SHFGPGPTLGKP 2300

```

Search completed: March 24, 2004, 13:17:10
Job time : 2.15608 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 13:14:29 ; Search time 5.22686 Seconds
(without alignments)
3567.110 Million cell updates/sec

Title: US-09-856-681A-4
Perfect score: 376
Sequence: 1 PPPAPQRVDSIQVHSSQPSG.....PPKPSFAPLSTSMKPNDACT 72

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match Length	DB	ID		Description

1	376	100.0	971	12	US-10-403-676-46	Sequence 46, Appl
2	376	100.0	971	15	US-10-449-548-46	Sequence 46, Appl
3	376	100.0	981	12	US-10-403-676-30	Sequence 30, Appl
4	376	100.0	981	15	US-10-449-548-30	Sequence 30, Appl
5	376	100.0	998	12	US-10-403-676-20	Sequence 20, Appl
6	376	100.0	998	15	US-10-449-548-20	Sequence 20, Appl
7	376	100.0	1018	12	US-10-403-676-28	Sequence 28, Appl
8	376	100.0	1018	15	US-10-449-548-28	Sequence 28, Appl
9	376	100.0	1030	12	US-10-016-248-63	Sequence 63, Appl
10	376	100.0	1035	12	US-10-403-676-18	Sequence 18, Appl
11	376	100.0	1035	15	US-10-449-548-18	Sequence 18, Appl
12	376	100.0	1047	10	US-09-957-187-85	Sequence 85, Appl
13	376	100.0	1047	12	US-10-403-676-14	Sequence 14, Appl
14	376	100.0	1047	12	US-10-403-676-48	Sequence 48, Appl
15	376	100.0	1047	15	US-10-449-548-14	Sequence 14, Appl
16	376	100.0	1047	15	US-10-449-548-48	Sequence 48, Appl
17	163.5	43.5	1011	12	US-10-016-248-61	Sequence 61, Appl
18	163.5	43.5	1022	12	US-10-016-248-60	Sequence 60, Appl
19	163.5	43.5	1032	12	US-10-016-248-16	Sequence 16, Appl
20	163.5	43.5	1035	12	US-10-016-248-12	Sequence 12, Appl
21	163.5	43.5	1088	12	US-10-016-248-20	Sequence 20, Appl
22	87	23.1	961	15	US-10-231-913-102	Sequence 102, App
23	86	22.9	873	9	US-09-843-245-3	Sequence 3, Appli
24	84	22.3	1448	16	US-10-389-566-786	Sequence 786, App
25	80.5	21.4	374	14	US-10-239-431A-20	Sequence 20, Appl
26	80.5	21.4	564	14	US-10-239-431A-8	Sequence 8, Appli
27	79	21.0	213	12	US-10-425-114-67112	Sequence 67112, A
28	79	21.0	493	12	US-10-424-599-153103	Sequence 153103,
29	77.5	20.6	693	15	US-10-108-260A-3849	Sequence 3849, Ap
30	76	20.2	174	12	US-10-425-114-68887	Sequence 68887, A
31	76	20.2	196	12	US-10-425-114-50652	Sequence 50652, A
32	76	20.2	196	12	US-10-425-114-69198	Sequence 69198, A
33	76	20.2	196	12	US-10-425-114-69259	Sequence 69259, A
34	76	20.2	196	12	US-10-425-114-70738	Sequence 70738, A
35	74.5	19.8	1389	12	US-10-282-122A-70151	Sequence 70151, A
36	74	19.7	628	14	US-10-213-990-36	Sequence 36, Appl
37	74	19.7	884	14	US-10-303-683-21	Sequence 21, Appl
38	74	19.7	894	10	US-09-291-417-18	Sequence 18, Appl
39	74	19.7	894	14	US-10-303-683-20	Sequence 20, Appl
40	73.5	19.5	2665	9	US-09-864-761-34248	Sequence 34248, A
41	73.5	19.5	3664	14	US-10-177-293-423	Sequence 423, App
42	73	19.4	200	14	US-10-259-165-300	Sequence 300, App
43	73	19.4	4019	9	US-09-738-973-425	Sequence 425, App
44	73	19.4	4019	9	US-09-854-133-425	Sequence 425, App
45	73	19.4	4019	14	US-10-144-649A-425	Sequence 425, App

ALIGNMENTS

RESULT 1

US-10-403-676-46

; Sequence 46, Application US/10403676

; Publication No. US20040029150A1

; GENERAL INFORMATION:

; APPLICANT: Alsobrook II, John

; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Grosse, William M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: LaRochelle, William J.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME,
AND METHODS OF USE
; FILE REFERENCE: 21402-573B
; CURRENT APPLICATION NUMBER: US/10/403,676
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 09/957,187
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 46

; LENGTH: 971
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-676-46

Query Match 100.0%; Score 376; DB 12; Length 971;
Best Local Similarity 100.0%; Pred. No. 3.2e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
|||||
Db 900 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 959

Qy 61 LSTSMKPNDACT 72
|||||
Db 960 LSTSMKPNDACT 971

RESULT 2

US-10-449-548-46

; Sequence 46, Application US/10449548
; Publication No. US20040018977A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez, Enrique
; APPLICANT: Anderson, David W.
; APPLICANT: Dhanabal, Mohanraj
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: LaRochelle, William J.
; APPLICANT: Li, Li
; APPLICANT: Lichenstein, Henri
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: SEMAPHORIN-LIKE PROTEINS AND METHODS OF USING SAME
; FILE REFERENCE: 15966-540CIP2
; CURRENT APPLICATION NUMBER: US/10/449,548
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 60/234,082
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/174,485
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 10/403,676
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/384,798
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/402,407
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/443,062


```
; PRIOR FILING DATE: 2003-01-28
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 46
; LENGTH: 971
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-449-548-46
```

Query Match 100.0%; Score 376; DB 15; Length 971;
Best Local Similarity 100.0%; Pred. No. 3.2e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy          1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
            |||
Db          900 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 959
            |||
Qy          61 LSTSMKPNDACT 72
            |||
Db          960 LSTSMKPNDACT 971

```

RESULT 3

US-10-403-676-30

; Sequence 30, Application US/10403676

; Publication No. US20040029150A1

; GENERAL INFORMATION:

;
; APPLICANT: Alsobrook II, John
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Grosse, William M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: LaRochelle, William J.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.

```

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME,
AND METHODS OF USE
; FILE REFERENCE: 21402-573B
; CURRENT APPLICATION NUMBER: US/10/403,676
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 09/957,187
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 30
; LENGTH: 981
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-676-30

```

```

Query Match          100.0%; Score 376; DB 12; Length 981;
Best Local Similarity 100.0%; Pred. No. 3.2e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      907 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 966

Qy      61 LSTSMKPNDACT 72
          |||||||||
Db      967 LSTSMKPNDACT 978

```

```

RESULT 4
US-10-449-548-30
; Sequence 30, Application US/10449548
; Publication No. US20040018977A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez, Enrique
; APPLICANT: Anderson, David W.
; APPLICANT: Dhanabal, Mohanraj
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: LaRochelle, William J.

```

```

; APPLICANT: Li, Li
; APPLICANT: Lichenstein, Henri
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: SEMAPHORIN-LIKE PROTEINS AND METHODS OF USING SAME
; FILE REFERENCE: 15966-540CIP2
; CURRENT APPLICATION NUMBER: US/10/449,548
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 60/234,082
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/174,485
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 10/403,676
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/384,798
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/402,407
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/443,062
; PRIOR FILING DATE: 2003-01-28
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 30
; LENGTH: 981
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-449-548-30

```

```

Query Match          100.0%; Score 376; DB 15; Length 981;
Best Local Similarity 100.0%; Pred. No. 3.2e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 PPPAPQQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
        |||
Db      907 PPPAPQQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 966
        |||

Qy      61 LSTSMKPNDACT 72
        |||
Db      967 LSTSMKPNDACT 978

```

```

RESULT 5
US-10-403-676-20
; Sequence 20, Application US/10403676
; Publication No. US20040029150A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John

```

; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Grosse, William M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: LaRochelle, William J.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME,
AND METHODS OF USE
; FILE REFERENCE: 21402-573B
; CURRENT APPLICATION NUMBER: US/10/403,676
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 09/957,187
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 20

; LENGTH: 998
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-676-20

Query Match 100.0%; Score 376; DB 12; Length 998;
Best Local Similarity 100.0%; Pred. No. 3.3e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
|
Db 924 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 983

Qy 61 LSTSMKPNDACT 72
|
Db 984 LSTSMKPNDACT 995

RESULT 6

US-10-449-548-20
; Sequence 20, Application US/10449548
; Publication No. US20040018977A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez, Enrique
; APPLICANT: Anderson, David W.
; APPLICANT: Dhanabal, Mohanraj
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: LaRochelle, William J.
; APPLICANT: Li, Li
; APPLICANT: Lichenstein, Henri
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: SEMAPHORIN-LIKE PROTEINS AND METHODS OF USING SAME
; FILE REFERENCE: 15966-540CIP2
; CURRENT APPLICATION NUMBER: US/10/449,548
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 60/234,082
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/174,485
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 10/403,676
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/384,798
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/402,407
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/443,062

```
; PRIOR FILING DATE: 2003-01-28
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 20
; LENGTH: 998
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-449-548-20
```

Query Match 100.0%; Score 376; DB 15; Length 998;
Best Local Similarity 100.0%; Pred. No. 3.3e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      924 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 983

Qy      61 LSTSMKPNDACT 72
      |||||||||||
Db      984 LSTSMKPNDACT 995

```

RESULT 7

US-10-403-676-28

; Sequence 28, Application US/10403676

; Publication No. US20040029150A1

; GENERAL INFORMATION:

; APPLICANT: Alsobrook II, John

; APPLICANT: Anderson, David W.

; APPLICANT: Boldog, Ferenc L.

; APPLICANT: Burgess, Catherine E.

; APPLICANT: Casman, Stacie J.

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Gerlach, Valerie L.

; APPLICANT: Grosse, William M.

; APPLICANT: Guo, Xiaojia

; APPLICANT: Gusev, Vladimir Y.

; APPLICANT: Ji, Weizhen

; APPLICANT: LaRochelle, William J.

; APPLICANT: Lepley, Denise M.

; APPLICANT: Li, Li

; APPLICANT: Liu, Xiaohong

; APPLICANT: MacDougall, John R.

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Millet, Isabelle

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Patturajan, Meera

; APPLICANT: Peyman, John A.

; APPLICANT: Rastelli, Luca

; APPLICANT: Reiger, Daniel

; APPLICANT: Rothenberg, Mark E.

; APPLICANT: Shimkets, Richard A.

; APPLICANT: Stone, David J.

; APPLICANT: Taupier, Raymond J.

; APPLICANT: Vernet, Corine

; APPLICANT: Zerhusen, Bryan D.

```

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME,
AND METHODS OF USE
; FILE REFERENCE: 21402-573B
; CURRENT APPLICATION NUMBER: US/10/403,676
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 09/957,187
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 28
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-676-28

```

```

Query Match          100.0%; Score 376; DB 12; Length 1018;
Best Local Similarity 100.0%; Pred. No. 3.4e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
        |||
Db      944 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 1003

Qy      61 LSTSMKPNDACT 72
        |||
Db      1004 LSTSMKPNDACT 1015

```

RESULT 8

```

US-10-449-548-28
; Sequence 28, Application US/10449548
; Publication No. US20040018977A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez, Enrique
; APPLICANT: Anderson, David W.
; APPLICANT: Dhanabal, Mohanraj
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: LaRochelle, William J.

```

```

; APPLICANT: Li, Li
; APPLICANT: Lichenstein, Henri
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: SEMAPHORIN-LIKE PROTEINS AND METHODS OF USING SAME
; FILE REFERENCE: 15966-540CIP2
; CURRENT APPLICATION NUMBER: US/10/449,548
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 60/234,082
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/174,485
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 10/403,676
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/384,798
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/402,407
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/443,062
; PRIOR FILING DATE: 2003-01-28
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 28
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-449-548-28

```

```

Query Match          100.0%; Score 376; DB 15; Length 1018;
Best Local Similarity 100.0%; Pred. No. 3.4e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      944 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 1003

Qy      61 LSTSMKPNDACT 72
        |||||||||||
Db      1004 LSTSMKPNDACT 1015

```

```

RESULT 9
US-10-016-248-63
; Sequence 63, Application US/10016248
; Publication No. US20040033491A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.

```



```

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-218
; CURRENT APPLICATION NUMBER: US/10/016,248
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/254,329
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/291,037
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/255,648
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 60/297,173
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/309,258
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/326,393
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/315,639
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 1030
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-016-248-63

```

```

Query Match          100.0%; Score 376; DB 12; Length 1030;
Best Local Similarity 100.0%; Pred. No. 3.4e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
          |||
Db      959 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 1018

Qy      61 LSTSMKPNDACT 72
          |||
Db      1019 LSTSMKPNDACT 1030

```

RESULT 10

```

US-10-403-676-18
; Sequence 18, Application US/10403676
; Publication No. US20040029150A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Grosse, William M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: LaRochelle, William J.
; APPLICANT: Lepley, Denise M.

```

```

; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME,
AND METHODS OF USE
; FILE REFERENCE: 21402-573B
; CURRENT APPLICATION NUMBER: US/10/403,676
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 09/957,187
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 18
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-676-18

```

```

Query Match          100.0%; Score 376; DB 12; Length 1035;
Best Local Similarity 100.0%; Pred. No. 3.4e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      . 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
          |||
Db      961 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 1020

```

Qy 61 LSTSMKPNDACT 72
 |||||
Db 1021 LSTSMKPNDACT 1032

RESULT 11
US-10-449-548-18
; Sequence 18, Application US/10449548
; Publication No. US20040018977A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez, Enrique
; APPLICANT: Anderson, David W.
; APPLICANT: Dhanabal, Mohanraj
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: LaRochelle, William J.
; APPLICANT: Li, Li
; APPLICANT: Lichenstein, Henri
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: SEMAPHORIN-LIKE PROTEINS AND METHODS OF USING SAME
; FILE REFERENCE: 15966-540CIP2
; CURRENT APPLICATION NUMBER: US/10/449,548
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 60/234,082
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/174,485
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 10/403,676
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/384,798
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/402,407
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/443,062
; PRIOR FILING DATE: 2003-01-28
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 18
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-449-548-18

Query Match 100.0%; Score 376; DB 15; Length 1035;
Best Local Similarity 100.0%; Pred. No. 3.4e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
|
Db 961 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 1020

Qy 61 LSTSMKPNDACT 72
|
Db 1021 LSTSMKPNDACT 1032

RESULT 12

US-09-957-187-85

; Sequence 85, Application US/09957187
; Publication No. US20030054514A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A.
; APPLICANT: LaRochelle, William
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
; FILE REFERENCE: 15966-540 CIP
; CURRENT APPLICATION NUMBER: US/09/957,187
; CURRENT FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/234,082
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/174,485
; PRIOR FILING DATE: 2000-01-04
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-957-187-85

Query Match 100.0%; Score 376; DB 10; Length 1047;
Best Local Similarity 100.0%; Pred. No. 3.5e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
|
Db 976 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 1035

Qy 61 LSTSMKPNDACT 72
|
Db 1036 LSTSMKPNDACT 1047

RESULT 13

US-10-403-676-14

; Sequence 14, Application US/10403676
; Publication No. US20040029150A1
; GENERAL INFORMATION:

; APPLICANT: Alsobrook II, John
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Grosse, William M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: LaRochelle, William J.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME,
AND METHODS OF USE
; FILE REFERENCE: 21402-573B
; CURRENT APPLICATION NUMBER: US/10/403,676
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 09/957,187
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1


```

; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 09/957,187
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 48
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-676-48

```

```

Query Match          100.0%; Score 376; DB 12; Length 1047;
Best Local Similarity 100.0%; Pred. No. 3.5e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
        |||
Db      976 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 1035

Qy      61 LSTSMKPNDACT 72
        |||
Db      1036 LSTSMKPNDACT 1047

```

RESULT 15

US-10-449-548-14

```

; Sequence 14, Application US/10449548
; Publication No. US20040018977A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez, Enrique
; APPLICANT: Anderson, David W.
; APPLICANT: Dhanabal, Mohanraj
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: LaRochelle, William J.
; APPLICANT: Li, Li
; APPLICANT: Lichenstein, Henri
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Padigaru, Muralidhara

```

```
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: SEMAPHORIN-LIKE PROTEINS AND METHODS OF USING SAME
; FILE REFERENCE: 15966-540CIP2
; CURRENT APPLICATION NUMBER: US/10/449,548
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 60/234,082
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/174,485
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 10/403,676
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/384,798
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/402,407
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/443,062
; PRIOR FILING DATE: 2003-01-28
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 14
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-449-548-14
```

```
Query Match          100.0%; Score 376; DB 15; Length 1047;
Best Local Similarity 100.0%; Pred. No. 3.5e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      976 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 1035

Qy      61 LSTSMKPNDACT 72
          |||||||||||
Db      1036 LSTSMKPNDACT 1047
```

```
Search completed: March 24, 2004, 13:19:32
Job time : 6.22686 secs
```


OM protein - protein search, using sw model

Run on: March 24, 2004, 13:11:03 ; Search time 5.03085 Seconds
(without alignments)
4515.598 Million cell updates/sec

Title: US-09-856-681A-4
Perfect score: 376
Sequence: 1 PPPAPQRVDSIQVHSSQPSG.....PPKPSFAPLSTSMKPNDACT 72

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description

1	376	100.0	507	4	Q96T04	Q96t04 homo sapien
2	376	100.0	562	4	Q96SY4	Q96sy4 homo sapien
3	376	100.0	562	4	Q8NC49	Q8nc49 homo sapien
4	376	100.0	574	4	Q96SM8	Q96sm8 homo sapien
5	376	100.0	699	4	Q96SW4	Q96sw4 homo sapien
6	376	100.0	1005	11	Q9EQ71	Q9eq71 mus musculu
7	169.5	45.1	1009	11	Q80TD0	Q80td0 mus musculu
8	163.5	43.5	416	6	Q95KA6	Q95ka6 macaca fasc
9	163.5	43.5	451	4	Q9H9K4	Q9h9k4 homo sapien
10	163.5	43.5	464	4	Q9H9G5	Q9h9g5 homo sapien
11	163.5	43.5	998	4	Q8Nfy6	Q8nfy6 homo sapien
12	163.5	43.5	1011	4	Q8Nfy3	Q8nfy3 homo sapien
13	163.5	43.5	1017	4	Q8Nfy5	Q8nfy5 homo sapien
14	163.5	43.5	1022	4	Q9P249	Q9p249 homo sapien
15	163.5	43.5	1073	4	Q8Nfy4	Q8nfy4 homo sapien
16	81	21.5	994	13	Q7ZZ40	Q7zz40 brachydanio
17	80.5	21.4	508	11	Q8CD55	Q8cd55 mus musculu
18	80.5	21.4	533	11	Q7TQE2	Q7tqe2 mus musculu
19	80.5	21.4	564	11	Q8CBM0	Q8cbm0 mus musculu
20	80.5	21.4	876	5	Q9XZN5	Q9xzn5 mya arenari
21	80.5	21.4	1322	11	Q9QXI0	Q9qxi0 rattus norv
22	80.5	21.4	1912	11	Q9ERC1	Q9erc1 rattus norv
23	80	21.3	477	6	O97600	O97600 oryctolagus
24	78	20.7	1220	5	Q9GPS9	Q9gps9 dictyosteli
25	77.5	20.6	616	4	Q9H6K5	Q9h6k5 homo sapien
26	77.5	20.6	1111	10	Q9SZL9	Q9szl9 arabidopsis
27	77	20.5	144	10	Q8GVN3	Q8gvn3 oryza sativ
28	77	20.5	175	10	Q9M1T6	Q9mlt6 arabidopsis
29	77	20.5	175	10	Q8GWV9	Q8gww9 arabidopsis
30	77	20.5	960	11	Q921L2	Q921l2 mus musculu
31	76.5	20.3	744	10	O65375	O65375 arabidopsis
32	76	20.2	508	4	Q9NXZ9	Q9nxz9 homo sapien
33	76	20.2	508	4	O76049	O76049 homo sapien
34	75.5	20.1	312	10	Q9SI74	Q9si74 arabidopsis
35	75.5	20.1	417	4	Q86VU4	Q86vu4 homo sapien
36	75.5	20.1	441	5	Q965J5	Q965j5 caenorhabdi
37	75.5	20.1	493	4	Q8NFN7	Q8nfn7 homo sapien
38	75.5	20.1	698	5	Q8MSL9	Q8msl9 drosophila
39	75.5	20.1	732	3	Q8J1Y5	Q8jly5 ashbya goss
40	75.5	20.1	735	5	Q9W3N8	Q9w3n8 drosophila
41	75.5	20.1	745	5	Q9U484	Q9u484 drosophila
42	75.5	20.1	745	5	Q9W164	Q9w164 drosophila
43	75.5	20.1	1216	3	Q9C276	Q9c276 neurospora
44	74.5	19.8	449	5	O46062	O46062 drosophila
45	74.5	19.8	653	5	Q8MR25	Q8mr25 drosophila

ALIGNMENTS

RESULT 1

Q96T04

ID Q96T04 PRELIMINARY; PRT; 507 AA.

AC Q96T04;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Hypothetical protein FLJ14533.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AK027439; BAB55111.1; -.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR003659; Plexin-like.
 DR SMART; SM00423; PSI; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 507 AA; 55464 MW; 8CC567B438C51B39 CRC64;

Query Match 100.0%; Score 376; DB 4; Length 507;
 Best Local Similarity 100.0%; Pred. No. 3.8e-33;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 436 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 495

 QY 61 LSTSMKPNDACT 72
 ||||||||||||
 Db 496 LSTSMKPNDACT 507

RESULT 2

Q96SY4

ID Q96SY4 PRELIMINARY; PRT; 562 AA.
 AC Q96SY4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein FLJ14565.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;

RESULT 4

Q96SM8

ID Q96SM8 PRELIMINARY; PRT; 574 AA.
 AC Q96SM8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein FLJ14748.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AK027654; BAB55269.1; -.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR003659; Plexin-like.
 DR SMART; SM00423; PSI; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 574 AA; 62822 MW; 0C79E01A4117A495 CRC64;

Query Match 100.0%; Score 376; DB 4; Length 574;
 Best Local Similarity 100.0%; Pred. No. 4.3e-33;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQVRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
 |||||
 Db 503 PPPAPQVRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 562
 Qy 61 LSTSMKPNDACT 72
 |||||
 Db 563 LSTSMKPNDACT 574

RESULT 5

Q96SW4

ID Q96SW4 PRELIMINARY; PRT; 699 AA.
 AC Q96SW4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein FLJ14595.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AK027501; BAB55158.1; -.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 699 AA; 76723 MW; 2E5F111D59741394 CRC64;

Query Match 100.0%; Score 376; DB 4; Length 699;
 Best Local Similarity 100.0%; Pred. No. 5.4e-33;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 628 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 687

 Qy 61 LSTSMKPNDACT 72
 ||||||||||||
 Db 688 LSTSMKPNDACT 699

RESULT 6

Q9EQ71

ID Q9EQ71 PRELIMINARY; PRT; 1005 AA.
 AC Q9EQ71;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Axon guidance signal SEMA6A1.
 GN SEMA6A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20564339; PubMed=10993894;
 RA Klostermann A., Lutz B., Gertler F., Behl C.;
 RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-
 RT 1/Sema6A-1) bind to the Enabled/Vasodilator-stimulated Phosphoprotein-
 RT like Protein (EVL) via a novel carboxyl-terminal Zyxin-like domain.";
 RL J. Biol. Chem. 275:39647-39653(2000).
 DR EMBL; AF288666; AAG29494.1; -.

DR MGD; MGI:1203727; Sema6a.
 DR GO; GO:0016021; C:integral to membrane; ISS.
 DR GO; GO:0008580; F:cytoskeletal regulator activity; ISS.
 DR GO; GO:0007411; P:axon guidance; ISS.
 DR GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; ISS.
 DR GO; GO:0007399; P:neurogenesis; ISS.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 SQ SEQUENCE 1005 AA; 111758 MW; 57B69927F45B079D CRC64;

Query Match 100.0%; Score 376; DB 11; Length 1005;
 Best Local Similarity 100.0%; Pred. No. 8.3e-33;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
 |||
 Db 934 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 993
 Qy 61 LSTSMKPNDACT 72
 |||
 Db 994 LSTSMKPNDACT 1005

RESULT 7

Q80TD0

ID Q80TD0 PRELIMINARY; PRT; 1009 AA.
 AC Q80TD0;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE MKIAA1479 protein (Fragment).
 GN MKIAA1479.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 DR EMBL; AK122515; BAC65797.1; -.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 1009 AA; 112808 MW; 7509F0B67332316B CRC64;

Query Match 43.5%; Score 163.5; DB 4; Length 464;
Best Local Similarity 50.7%; Pred. No. 9.4e-10;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

RESULT 11

```

ID      Q8NFY6          PRELIMINARY;          PRT;      998 AA.
AC      Q8NFY6;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Semaphorin 6D isoform 2.
GN      SEMA6D.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RA      Qu X., Zhai Y., Wei H., Yu Y., Tang F., He F.;
RT      "Homo sapiens semaphorin 6D isoform 2 (SEMA6D.2) mRNA, complete cds.";
RL      Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR      EMBL; AF389427; AAM69450.1; -.
DR      InterPro; IPR001627; Sema.
DR      Pfam; PF01403; Sema; 1.
DR      SMART; SM00630; Sema; 1.
SQ      SEQUENCE      998 AA;  111730 MW;  3F46D6872E8D5344 CRC64;

```

Query Match 43.5%; Score 163.5; DB 4; Length 998;
Best Local Similarity 50.7%; Pred. No. 2.3e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

RESULT 12

ID Q8NFY3 PRELIMINARY; PRT; 1011 AA.
AC Q8NFY3;

Search completed: March 24, 2004, 13:16:26
Job time : 6.03085 secs

OM protein - protein search, using sw model

Run on: March 24, 2004, 13:07:38 ; Search time 1.43739 Seconds
 (without alignments)
 2608.241 Million cell updates/sec

Title: US-09-856-681A-4
 Perfect score: 376
 Sequence: 1 PPPAPQVRVDSIQVHSSQPSG.....PPKPSFAPLSTSMKPNDACT 72

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				
1	376	100.0	1030	1	SM6A_HUMAN	Q9h2e6 homo sapien
2	87	23.1	961	1	FGD1_HUMAN	P98174 homo sapien
3	86.5	23.0	564	1	ZYX_MOUSE	Q62523 mus musculu
4	86	22.9	862	1	M4K3_RAT	Q924i2 rattus norv
5	77	20.5	960	1	FGD1_MOUSE	P52734 mus musculu
6	75.5	20.1	397	1	GAT5_HUMAN	Q9bwx5 homo sapien
7	75.5	20.1	477	1	MAZ_HUMAN	P56270 homo sapien
8	75.5	20.1	5147	1	PCLO_HUMAN	Q9y6v0 homo sapien
9	74	19.7	452	1	HIS7_PHYPR	P28624 phytophthor
10	74	19.7	894	1	M4K3_HUMAN	Q8ivh8 homo sapien
11	73.5	19.5	3664	1	MINT_HUMAN	Q96t58 homo sapien
12	73.5	19.5	5085	1	PCLO_RAT	Q9jks6 rattus norv
13	73	19.4	4911	1	MLL3_HUMAN	Q8nez4 homo sapien
14	72.5	19.3	628	1	V70K_TYMV	P10357 turnip yell
15	72.5	19.3	5262	1	MLL2_HUMAN	O14686 homo sapien
16	71.5	19.0	446	1	TFE3_MOUSE	Q64092 mus musculu
17	71.5	19.0	668	1	SCEL_HUMAN	O95171 homo sapien

18	71.5	19.0	1004	1	PHC1_HUMAN	P78364	homo sapien
19	71.5	19.0	1522	1	PST1_SCHPO	Q09750	schizosacch
20	71	18.9	428	1	ELK1_HUMAN	P19419	homo sapien
21	71	18.9	621	1	APS_MOUSE	Q9jid9	mus musculu
22	70.5	18.8	344	1	ZIPA_SHEON	Q8ed69	shewanella
23	70	18.6	719	1	PRH1_SCHPO	Q03319	schizosacch
24	70	18.6	1012	1	PHC1_MOUSE	Q64028	mus musculu
25	69.5	18.5	331	1	MAZ_MESAU	P56670	mesocricetu
26	69.5	18.5	477	1	MAZ_MOUSE	P56671	mus musculu
27	69.5	18.5	1125	1	MAP4_MOUSE	P27546	mus musculu
28	68.5	18.2	625	1	R101_YEAST	P33400	saccharomyc
29	68.5	18.2	743	1	TFE3_HUMAN	P19532	homo sapien
30	68	18.1	737	1	SKN1_CANAL	P87024	candida alb
31	68	18.1	812	1	NAH2_HUMAN	Q9uby0	homo sapien
32	68	18.1	1259	1	AUT2_HUMAN	Q8wxx7	homo sapien
33	68	18.1	1273	1	SN3A_HUMAN	Q96st3	homo sapien
34	68	18.1	1282	1	SN3A_MOUSE	Q60520	mus musculu
35	67.5	18.0	429	1	ELK1_MOUSE	P41969	mus musculu
36	67.5	18.0	525	1	CO2A_HUMAN	Q92828	homo sapien
37	67.5	18.0	628	1	V70K_TYMVA	P20131	turnip yell
38	67.5	18.0	628	1	V70K_TYMCV	P28478	turnip yell
39	67.5	18.0	5179	1	MUC2_HUMAN	Q02817	homo sapien
40	67	17.8	315	1	YK04_CAEEL	P34292	caenorhabdi
41	67	17.8	529	1	DNB2_ADE05	P03265	human adeno
42	67	17.8	638	1	KNC0_YEAST	P53974	saccharomyc
43	67	17.8	779	1	SRP_DROME	P52172	drosophila
44	67	17.8	813	1	NAH2_RAT	P48763	rattus norv
45	67	17.8	1078	1	S24A_HUMAN	O95486	homo sapien

ALIGNMENTS

RESULT 1

SM6A_HUMAN

ID SM6A_HUMAN STANDARD; PRT; 1030 AA.

AC Q9H2E6; Q9P2H9;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Semaphorin 6A precursor (Semaphorin VIA) (Sema VIA) (Semaphorin 6A-1)
(SEMA6A-1).

GN SEMA6A OR KIAA1368.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND INTERACTION WITH EVL.

RX MEDLINE=20564339; PubMed=10993894;

RA Klostermann A., Lutz B., Gertler F., Behl C.;

RT "The orthologous human and murine semaphorin 6A-1 proteins

RT (SEMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated

RT phosphoprotein-like protein (EVL) via a novel carboxyl-terminal

RT zyxin-like domain.";

RL J. Biol. Chem. 275:39647-39653(2000).

RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XVI.
 RT The complete sequences of 150 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:65-73(2000).
 CC -!- FUNCTION: Can act as repulsive axon guidance cues. May play a role
 CC in channeling sympathetic axons into the sympathetic chains and
 CC controlling the temporal sequence of sympathetic target
 CC innervation (By similarity).
 CC -!- SUBUNIT: Active as a homodimer or oligomer. Interacts with EVL.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9H2E6-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9H2E6-2; Sequence=VSP_007113;
 CC Note=No experimental confirmation available;
 CC -!- SIMILARITY: Belongs to the semaphorin family.
 CC -!- SIMILARITY: Contains 1 Sema domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; AF279656; AAG29378.1; -.
 DR EMBL; AB037789; BAA92606.1; ALT_INIT.
 DR Genew; HGNC:10738; SEMA6A.
 DR MIM; 605885; -.
 DR GO; GO:0030424; C:axon; NAS.
 DR GO; GO:0016021; C:integral to membrane; NAS.
 DR GO; GO:0008580; F:cytoskeletal regulator activity; NAS.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0006915; P:apoptosis; NAS.
 DR GO; GO:0007411; P:axon guidance; NAS.
 DR GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; NAS.
 DR GO; GO:0007399; P:neurogenesis; NAS.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 KW Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 KW Developmental protein; Alternative splicing.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 1030 SEMAPHORIN 6A.
 FT DOMAIN 19 649 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 650 670 POTENTIAL.
 FT DOMAIN 671 1030 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 56 491 SEMA.

FT DOMAIN 792 819 PRO-RICH.
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 576 576 N -> NDISTPLPDNEMSYNTVY (in isoform 2).
 FT /FTid=VSP_007113.
 SQ SEQUENCE 1030 AA; 114368 MW; A57B79C10AEC4B34 CRC64;

Query Match 100.0%; Score 376; DB 1; Length 1030;
 Best Local Similarity 100.0%; Pred. No. 1.5e-28;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 959 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 1018

 Qy 61 LSTSMKPNDACT 72
 ||||||||||||
 Db 1019 LSTSMKPNDACT 1030

RESULT 2

FGD1_HUMAN

ID FGD1_HUMAN STANDARD; PRT; 961 AA.
 AC P98174; Q8N4D9;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Putative Rho/Rac guanine nucleotide exchange factor (Rho/Rac GEF)
 DE (Faciogenital dysplasia protein).
 GN FGD1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Craniofacial;
 RX MEDLINE=95042764; PubMed=7954831;
 RA Pasteris N.G., Cadle A., Logie L.J., Porteous M.E.M., Schwartz C.E.,
 RA Stevenson R.E., Glover T.W., Wilroy R.S., Gorski J.L.;
 RT "Isolation and characterization of the faciogenital dysplasia
 RT (Aarskog-Scott syndrome) gene: a putative Rho/Rac guanine nucleotide
 RT exchange factor.";
 RL Cell 79:669-678(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP VARIANT AAS HIS-522.
 RX MEDLINE=20546218; PubMed=11093277;
 RA Schwartz C.E., Gillesen-Kaesbach G., May M., Cappa M., Gorski J.L.,
 RA Steindl K., Neri G.;
 RT "Two novel mutations confirm FGD1 is responsible for the Aarskog
 RT syndrome.";
 RL Eur. J. Hum. Genet. 8:869-874(2000).
 RN [4]
 RP VARIANT AAS GLN-610.
 RX MEDLINE=20389563; PubMed=10930571;
 RA Orrico A., Galli L., Falciani M., Bracci M., Cavaliere M.L.,
 RA Rinaldi M.M., Musacchio A., Sorrentino V.;
 RT "A mutation in the pleckstrin homology (PH) domain of the FGD1 gene in
 RT an Italian family with faciogenital dysplasia (Aarskog-Scott
 RT syndrome).";
 RL FEBS Lett. 478:216-220(2000).
 CC -!- FUNCTION: ACTIVATES THE RAS-LIKE FAMILY OF RHO- AND RAC PROTEINS
 CC BY EXCHANGING BOUND GDP FOR FREE GTP.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in fetal heart, brain, lung, kidney
 CC and placenta. Less expressed in liver; adult heart, brain, lung,
 CC pancreas and skeletal muscle.
 CC -!- DISEASE: Defects in FGD1 are the cause of Aarskog-Scott syndrome
 CC (AAS) [MIM:305400]. This faciogenital dysplasia is a rare
 CC multisystemic disorder characterized by disproportionately short
 CC stature, and by facial, skeletal, and urogenital anomalies.
 CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
 CC -!- SIMILARITY: Contains 1 FYVE-type zinc finger.
 CC -!- SIMILARITY: Contains 2 PH domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U11690; AAA57004.1; -.
 DR EMBL; BC034530; AAH34530.1; -.
 DR HSSP; Q07889; 1AWE.

DR Genew; HGNC:3663; FGD1.
 DR MIM; 305400; -.
 DR GO; GO:0005085; F:guanyl-nucleotide exchange factor activity; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR001331; GDS_CDC24.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000219; RhoGEF.
 DR InterPro; IPR000306; Znf_FYVE.
 DR Pfam; PF01363; FYVE; 1.
 DR Pfam; PF00169; PH; 2.
 DR Pfam; PF00621; RhoGEF; 1.
 DR SMART; SM00064; FYVE; 1.
 DR SMART; SM00233; PH; 2.
 DR SMART; SM00325; RhoGEF; 1.
 DR PROSITE; PS00741; DH_1; FALSE_NEG.
 DR PROSITE; PS50010; DH_2; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 2.
 DR PROSITE; PS50178; ZF_FYVE; 1.
 KW Guanine-nucleotide releasing factor; Zinc-finger; Repeat;
 KW Disease mutation.
 FT DOMAIN 373 561 DH.
 FT DOMAIN 7 330 PRO-RICH.
 FT SITE 171 179 SH3-BINDING (POTENTIAL).
 FT SITE 179 187 SH3-BINDING (POTENTIAL).
 FT DOMAIN 590 689 PH 1.
 FT ZN_FING 730 790 FYVE-TYPE.
 FT DOMAIN 821 921 PH 2.
 FT VARIANT 522 522 R -> H (IN AAS).
 FT /FTid=VAR_015236.
 FT VARIANT 610 610 R -> Q (IN AAS).
 FT /FTid=VAR_015237.
 FT CONFLICT 10 23 AGPSEPEHPATNPP -> RRAFGARTPGHEPA (IN REF.
 FT 1).
 FT CONFLICT 195 195 A -> G (IN REF. 1).
 SQ SEQUENCE 961 AA; 106560 MW; 30963F7B9931E45C CRC64;

Query Match 23.1%; Score 87; DB 1; Length 961;
 Best Local Similarity 34.8%; Pred. No. 0.7;
 Matches 23; Conservative 5; Mismatches 22; Indels 16; Gaps 2;

Qy 2 PPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 61
 | |||: | | |::| | || | || |||||: :
 Db 127 PEGPQRL-----RSDPGPPTETPSQRP-----SPLKRAPGPKPQVPPKPSYLQM 170

 Qy 62 STSMKP 67
 |
 Db 171 PRMPPP 176

RESULT 3

ZYX_MOUSE

ID ZYX_MOUSE STANDARD; PRT; 564 AA.

AC Q62523; P70461;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Zyxin.
 GN ZYX.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR X Swiss Webster;
 RX MEDLINE=97094926; PubMed=8940160;
 RA Macalma T., Otte J., Hensler M.E., Bockholt S.M., Louis H.A.,
 RA Kalff-Suske M., Grzeschik K.H., von der Ahe D., Beckerle M.C.;
 RT "Molecular characterization of human zyxin."
 RL J. Biol. Chem. 271:31470-31478(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Otte J., Heischmann A., Breier G., Beckerle M.C., von der Ahe D.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: Adhesion plaque protein. Binds alpha-actinin and the CRP
 CC protein. May be a component of a signal transduction pathway that
 CC mediates adhesion-stimulated changes in gene expression (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; associates with the actin
 CC cytoskeleton near the adhesion plaques.
 CC -!- SIMILARITY: Contains 3 LIM zinc-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y07711; CAA68984.1; -.
 DR EMBL; X99063; CAA67510.1; -.
 DR MGD; MGI:103072; Zyx.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 3.
 DR ProDom; PD000094; LIM; 3.
 DR SMART; SM00132; LIM; 3.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE; PS50023; LIM_DOMAIN_2; 3.
 KW Repeat; LIM domain; Metal-binding; Zinc; Cell adhesion.
 FT DOMAIN 64 77 PRO-RICH.
 FT DOMAIN 94 138 PRO-RICH.
 FT DOMAIN 376 435 LIM 1.
 FT DOMAIN 436 495 LIM 2.
 FT DOMAIN 496 562 LIM 3.
 FT CONFLICT 215 215 R -> A (IN REF. 1).
 FT CONFLICT 284 292 IKKWCLRMP -> NQKMVPPDA (IN REF. 1).
 FT CONFLICT 484 484 S -> C (IN REF. 1).
 SQ SEQUENCE 564 AA; 60790 MW; 001E1B3C82ADA1EB CRC64;

Query Match

23.0%; Score 86.5; DB 1; Length 564;

Best Local Similarity 34.3%; Pred. No. 0.43;
Matches 23; Conservative 11; Mismatches 24; Indels 9; Gaps 4;

```
QY      1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
      ||| ||| :|:| || : | || :: |: | | : | : |||
Db      209 PPPQPQRKPQVQLH-VQPQAKP-HVQPQP-VSSANTQPRGPLSQAPT-----PAPKEFAP 259

QY      61 LSTSMKP 67
      :: |
Db      260 VAPKFTP 266
```

RESULT 4

M4K3_RAT

```
ID  M4K3_RAT      STANDARD;      PRT;      862 AA.
AC  Q924I2;
DT  10-OCT-2003 (Rel. 42, Created)
DT  10-OCT-2003 (Rel. 42, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  Mitogen-activated protein kinase kinase kinase kinase 3 (EC 2.7.1.37)
DE  (MAPK/ERK kinase kinase kinase 3) (MEK kinase kinase 3) (MEKKK 3)
DE  (Germinal center kinase related protein kinase) (GLK) (Fragment).
GN  MAP4K3.
OS  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=10116;
RN  [1]
RP  SEQUENCE FROM N.A., AND INTERACTION WITH SH3GL2.
RX  MEDLINE=21369947; PubMed=11384986;
RA  Ramjaun A.R., Angers A., Legendre-Guillemain V., Tong X.-K.,
RA  McPherson P.S.;
RT  "Endophilin regulates JNK activation through its interaction with the
RT  germinal center kinase-like kinase.";
RL  J. Biol. Chem. 276:28913-28919(2001).
CC  -!- FUNCTION: May play a role in the response to environmental stress.
CC  Appears to act upstream of the c-jun N-terminal pathway (By
CC  similarity).
CC  -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC  -!- COFACTOR: Magnesium (By similarity).
CC  -!- SUBUNIT: Interacts with SH3GL2. Interaction appears to regulate
CC  MAP4K3-mediated JNK activation.
CC  -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC  STE20 subfamily.
CC  -!- SIMILARITY: Contains 1 CNH domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF312224; AAK53214.1; -.
DR  HSSP; P24941; 1BUH.
DR  GO; GO:0005524; F:ATP binding; ISS.
```

DR GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; ISS.
 DR GO; GO:0007243; P:protein kinase cascade; ISS.
 DR GO; GO:0006950; P:response to stress; ISS.
 DR InterPro; IPR001180; Citron.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 KW ATP-binding; Transferase; Serine/threonine-protein kinase.
 FT NON_TER 1 1
 FT DOMAIN 5 262 PROTEIN_KINASE.
 FT DOMAIN 530 842 CNH.
 FT NP_BIND 11 19 ATP (BY SIMILARITY).
 FT BINDING 34 34 ATP (BY SIMILARITY).
 FT ACT_SITE 125 125 BY SIMILARITY.
 SQ SEQUENCE 862 AA; 97390 MW; 58013AC3B0A3287F CRC64;

Query Match 22.9%; Score 86; DB 1; Length 862;
 Best Local Similarity 34.7%; Pred. No. 0.77;
 Matches 25; Conservative 6; Mismatches 27; Indels 14; Gaps 3;

Qy 1 PPPAPQRVDSIQV----HSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKP 56
 ||| | : || : ||| : | | | : | | | || || | | |
 Db 400 PPPLPPPKPSISIPQDTHSSEDSNQG-TIKRCPS-----SGSPAKPSHVPPRPPPP 449
 Qy 57 SFAPLSTSMKPN 68
 | :: |
 Db 450 RLPPQKPAVLGN 461

RESULT 5

FGD1_MOUSE

ID FGD1_MOUSE STANDARD; PRT; 960 AA.
 AC P52734;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative Rho/Rac guanine nucleotide exchange factor (Rho/Rac GEF)
 DE (Faciogenital dysplasia protein homolog).
 GN FGD1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96081343; PubMed=8535076;
 RA Pasteris N.G., de Gouyon B., Cadle A.B., Campbell K., Herman G.E.,
 RA Gorski J.L.;

RT "Cloning and regional localization of the mouse faciogenital
RT dysplasia (Fgdl) gene.";
RL Mamm. Genome 6:658-661(1995).
CC -!- FUNCTION: ACTIVATES THE RAS-LIKE FAMILY OF RHO- AND RAC PROTEINS
CC BY EXCHANGING BOUND GDP FOR FREE GTP.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -!- SIMILARITY: Contains 1 FYVE-type zinc finger.
CC -!- SIMILARITY: Contains 2 PH domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U22325; AAA96001.1; -.
DR HSSP; Q07889; 1AWE.
DR MGD; MGI:104566; Fgdl.
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhoGEF.
DR InterPro; IPR000306; Znf_FYVE.
DR Pfam; PF01363; FYVE; 1.
DR Pfam; PF00169; PH; 2.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00064; FYVE; 1.
DR SMART; SM00233; PH; 2.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS00741; DH_1; FALSE_NEG.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 2.
DR PROSITE; PS50178; ZF_FYVE; 1.
KW Guanine-nucleotide releasing factor; Zinc-finger; Repeat.
FT DOMAIN 372 560 DH.
FT DOMAIN 7 330 PRO-RICH.
FT SITE 171 179 SH3-BINDING (POTENTIAL).
FT SITE 179 187 SH3-BINDING (POTENTIAL).
FT DOMAIN 589 688 PH 1.
FT ZN_FING 729 789 FYVE-TYPE.
FT DOMAIN 820 920 PH 2.
SQ SEQUENCE 960 AA; 106477 MW; 41C1B84DE490FC51 CRC64;

Query Match 20.5%; Score 77; DB 1; Length 960;
Best Local Similarity 31.8%; Pred. No. 6.3;
Matches 21; Conservative 5; Mismatches 24; Indels 16; Gaps 2;

Qy 2 PPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 61
| |||: | | :| | || | || |||||: :
Db 127 PEGPQRL-----RSDPGPTEIPGPRP-----SPLKRAPGPKPQVPPKPSYLQM 170
Qy 62 STSMKP 67
: |
Db 171 PRVLPP 176

RESULT 6

GAT5_HUMAN

ID GAT5_HUMAN STANDARD; PRT; 397 AA.
AC Q9BW5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Transcription factor GATA-5 (GATA binding factor-5).
GN GATA5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehvaeslaiho M.H., Liversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -!- FUNCTION: Binds to the functionally important CEF-1 nuclear
CC protein binding site in the cardiac-specific slow/cardiac troponin
CC C transcriptional enhancer. May play an important role in the
CC transcriptional program(s) that underlies smooth muscle cell
CC diversity (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Contains 2 GATA-type zinc fingers.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL499627; CAC36001.1; -.
 DR HSSP; P17679; 1GNF.
 DR Genew; HGNC:15802; GATA5.
 DR InterPro; IPR008013; GATA-N.
 DR InterPro; IPR000679; Znf_GATA.
 DR Pfam; PF05349; GATA-N; 1.
 DR Pfam; PF00320; GATA; 2.
 DR PRINTS; PR00619; GATAZNFINGER.
 DR SMART; SM00401; Znf_GATA; 2.
 DR PROSITE; PS00344; GATA_ZN_FINGER_1; 2.
 DR PROSITE; PS50114; GATA_ZN_FINGER_2; 2.
 KW Transcription regulation; Activator; DNA-binding; Zinc-finger;
 KW Nuclear protein.
 FT ZN_FING 189 213 GATA-TYPE.
 FT ZN_FING 243 267 GATA-TYPE.
 SQ SEQUENCE 397 AA; 41299 MW; 5DFBA02085695C57 CRC64;

Query Match 20.1%; Score 75.5; DB 1; Length 397;
 Best Local Similarity 29.3%; Pred. No. 3.3;
 Matches 22; Conservative 8; Mismatches 36; Indels 9; Gaps 1;

Qy 1 PPPAPQRVDSIQVHSSQP-----SGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKP 51
 | | : : || : | | | | : : | |||
 Db 277 PRPLAMKKESIQTRKRKPKTIKARGSSGSTRNASASPSAVASTDSSAATSKAKPSLASP 336
 Qy 52 VPPKPSFAPLSTSMK 66
 | | || || : :
 Db 337 VCPGPSMAPQASGQE 351

RESULT 7

MAZ_HUMAN

ID MAZ_HUMAN STANDARD; PRT; 477 AA.
 AC P56270; Q15703; Q99443;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Myc-associated zinc finger protein (MAZI) (Purine-binding
 DE transcription factor) (Pur-1) (ZF87) (ZIF87).
 GN MAZ.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92366479; PubMed=1502157;
 RA Bossone S.A., Asselin C., Patel A.J., Marcu K.B.;
 RT "MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences
 RT regulating transcriptional initiation and termination."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7452-7456(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Carcinoma;

RX MEDLINE=92232709; PubMed=1567856;
 RA Pyrc J.J., Moberg K.H., Hall D.J.;
 RT "Isolation of a novel cDNA encoding a zinc-finger protein that binds
 RT to two sites within the c-myc promoter.";
 RL Biochemistry 31:4102-4110(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreatic islets;
 RX MEDLINE=96428591; PubMed=8831693;
 RA Tsutsui H., Sakatsume O., Itakura K., Yokoyama K.K.;
 RT "Members of the MAZ family: a novel cDNA clone for MAZ from human
 RT pancreatic islet cells.";
 RL Biochem. Biophys. Res. Commun. 226:801-809(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96224025; PubMed=8626793;
 RA Parks C.L., Shenk T.;
 RT "The serotonin 1a receptor gene contains a TATA-less promoter that
 RT responds to MAZ and Sp1.";
 RL J. Biol. Chem. 271:4417-4430(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymphoblastoma;
 RX MEDLINE=98352105; PubMed=9685418;
 RA Song J., Murakami H., Tsutsui H., Tang X., Matsumura M., Itakura K.,
 RA Kanazawa I., Sun K., Yokoyama K.K.;
 RT "Genomic organization and expression of a human gene for Myc-
 RT associated zinc finger protein (MAZ).";
 RL J. Biol. Chem. 273:20603-20614(1998).
 CC -!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR WITH DUAL ROLES
 CC IN TRANSCRIPTION INITIATION AND TERMINATION. BINDS TO TWO SITES,
 CC ME1A1 AND ME1A2, WITHIN THE C-MYC PROMOTER HAVING GREATER
 CC AFFINITY FOR THE FORMER. ALSO BINDS TO MULTIPLE G/C-RICH SITES
 CC WITHIN THE PROMOTER OF THE SP1 FAMILY OF TRANSCRIPTION FACTORS.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal
 CC muscle and pancreas. Seems not to be expressed in kidney.
 CC -!- SIMILARITY: Contains 6 C2H2-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M94046; -; NOT_ANNOTATED_CDS.
 DR EMBL; M93339; -; NOT_ANNOTATED_CDS.
 DR EMBL; D85131; BAA12728.1; ALT_INIT.
 DR EMBL; U33819; AAB04121.1; ALT_INIT.
 DR EMBL; AB017335; BAA33064.1; -.
 DR PIR; A42170; A42170.
 DR TRANSFAC; T00490; -.
 DR TRANSFAC; T02305; -.
 DR Genew; HGNC:6914; MAZ.
 DR MIM; 600999; -.

DR GO; GO:0006367; P:transcription initiation from Pol II promoter; TAS.
 DR GO; GO:0006369; P:transcription termination from Pol II promoter; TAS.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 6.
 DR ProDom; PD000003; Znf_C2H2; 1.
 DR SMART; SM00355; Znf_C2H2; 6.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 5.
 KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;
 KW RNA-binding; Repeat; Nuclear protein.
 FT ZN_FING 190 212 C2H2-TYPE 1.
 FT ZN_FING 279 301 C2H2-TYPE 2.
 FT ZN_FING 307 329 C2H2-TYPE 3.
 FT ZN_FING 337 360 C2H2-TYPE 4.
 FT ZN_FING 366 388 C2H2-TYPE 5.
 FT ZN_FING 392 413 C2H2-TYPE 6 (ATYPICAL).
 FT DOMAIN 96 108 POLY-ALA.
 FT DOMAIN 133 139 POLY-PRO.
 FT DOMAIN 157 161 POLY-ALA.
 FT DOMAIN 245 249 POLY-GLY.
 FT DOMAIN 435 449 POLY-ALA.
 FT CONFLICT 259 259 MISSING (IN REF. 3).
 FT CONFLICT 401 401 L -> M (IN REF. 2 AND 4).
 FT CONFLICT 443 447 MISSING (IN REF. 3).
 SQ SEQUENCE 477 AA; 48607 MW; C04C80F32C3C6825 CRC64;

Query Match 20.1%; Score 75.5; DB 1; Length 477;
 Best Local Similarity 25.8%; Pred. No. 4;
 Matches 23; Conservative 15; Mismatches 32; Indels 19; Gaps 3;

Qy 1 PPPAPQ-----RVDSIQV-----HSSQPSGQAVTVSRQPSLNAYNSLTRSGLK 43
 ||| || :|| :| :: : | : |: | :: : ||
 Db 69 PPPTPQAPAAEPLQVDLLFVLAAAQESAAAAAAAAAAAAVAAAPPAPAAASTVDTAALK 128
 Qy 44 RTPSLKPDVPPKPSFAPLSTSMKPNDACT 72
 : |: | || | || : : | ||
 Db 129 QPPA--PPPPPPVSAPEAAAPPASAAT 155

RESULT 8

PCLO_HUMAN

ID PCLO_HUMAN STANDARD; PRT; 5147 AA.
 AC Q9Y6V0; O43373; O60305; Q9BVC8; Q9UIV2; Q9Y6U9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Piccolo protein (Aczonin) (Fragments).
 GN PCLO OR ACZ OR KIAA0559.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-759 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99439764; PubMed=10508862;
 RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,

RA Kilimann M.W.;
 RT "Aczonin, a 550-kd putative scaffolding protein of presynaptic active
 RT zones, shares homology regions with rim and bassoon and binds
 RT profilin.";
 RL J. Cell Biol. 147:151-162(1999).
 RN [2]
 RP SEQUENCE OF 552-4404 FROM N.A.
 RA Kraemer J., Wollam C., Wohldmann P., McGrane B.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE OF 3619-5147 FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 RN [4]
 RP SEQUENCE OF 4405-4439 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 4405-5147 FROM N.A.
 RA Kalicki J., Elliott G.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: May act as a scaffolding protein involved in the
 CC organization of synaptic active zones and in synaptic vesicle
 CC trafficking (By similarity).
 CC -!- SUBUNIT: Interacts with Rabac1/Pral, RIMS2 and profilin (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
 CC synaptic junctions (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;

```

CC      Name=1;
CC      IsoId=Q9Y6V0-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=Q9Y6V0-2; Sequence=VSP_003923, VSP_003924, VSP_003925,
CC      VSP_003926, VSP_003927;
CC      Note=No experimental confirmation available;
CC      -!- DOMAIN: C2 domain 1 is involved in binding calcium and
CC      phospholipids. Calcium binds with low affinity but with high
CC      specificity and induces a large conformational change.
CC      -!- SIMILARITY: Contains 2 C2 domains.
CC      -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Y19188; CAB60727.1; -.
DR      EMBL; AC004903; AAD20936.1; -.
DR      EMBL; AC004886; AAD21789.1; -.
DR      EMBL; AB011131; BAA25485.1; -.
DR      EMBL; BC001304; AAH01304.1; -.
DR      EMBL; AC004082; AAB97937.1; -.
DR      PIR; T00634; T00634.
DR      HSSP; P04410; 1A25.
DR      Genew; HGNC:13406; PCLO.
DR      MIM; 604918; -.
DR      GO; GO:0005856; C:cytoskeleton; NAS.
DR      GO; GO:0045202; C:synaptic junction; ISS.
DR      GO; GO:0005509; F:calcium ion binding; ISS.
DR      GO; GO:0005544; F:calcium-dependent phospholipid binding; ISS.
DR      GO; GO:0005522; F:profilin binding; ISS.
DR      GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
DR      GO; GO:0016080; P:synaptic vesicle targeting; ISS.
DR      InterPro; IPR000008; C2.
DR      InterPro; IPR001565; Synaptotagmin.
DR      PRINTS; PR00360; C2DOMAIN.
DR      PRINTS; PR00399; SYNAPTOTAGMN.
DR      SMART; SM00239; C2; 2.
DR      PROSITE; PS00499; C2_DOMAIN_1; 1.
DR      PROSITE; PS50004; C2_DOMAIN_2; 2.
KW      Calcium/phospholipid-binding; Zinc; Metal-binding; Zinc-finger;
KW      Repeat; Alternative splicing.
FT      NON_TER      1      1
FT      DOMAIN      400      465      10 X 10 AA TANDEM APPROXIMATE REPEATS OF
FT      P-A-K-P-Q-P-Q-Q-P-X.
FT      ZN_FING      499      523      C4-TYPE (POTENTIAL).
FT      ZN_FING      969      992      C4-TYPE (POTENTIAL).
FT      NON_CONS     1010     1011
FT      DOMAIN      2300     2325      POLY-PRO.
FT      DOMAIN      4391     4442      PDZ.
FT      DOMAIN      4544     4633      C2 DOMAIN 1.
FT      DOMAIN      5031     5121      C2 DOMAIN 2.
FT      VARSPLIC     4404     4404      S -> SGNGLGIRIVGGKEIPGHSGEIGAYIAKILPGGSAAE

```

FT QTGKLMEG (in isoform 2).
 FT /FTId=VSP_003923.
 FT VARSPLIC 4534 4534 K -> KPTDGTKVVSHPIGEIQ (in isoform 2).
 FT /FTId=VSP_003924.
 FT VARSPLIC 4576 4576 G -> GQVMVVQNAS (in isoform 2).
 FT /FTId=VSP_003925.
 FT VARSPLIC 4757 4761 TAHKS -> SKRRK (in isoform 2).
 FT /FTId=VSP_003926.
 FT VARSPLIC 4762 5147 Missing (in isoform 2).
 FT /FTId=VSP_003927.
 SQ SEQUENCE 5147 AA; 563537 MW; CD5D84990498CD3C CRC64;

Query Match 20.1%; Score 75.5; DB 1; Length 5147;
 Best Local Similarity 31.4%; Pred. No. 59;
 Matches 22; Conservative 5; Mismatches 18; Indels 25; Gaps 3;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
 ||| | : || || | : || || : ||
 Db 2378 PPPVPPKPSSI-----PSGLVFTHRPEPS-----KPPIAPKPVIPQ 2413
 Qy 61 L-STSMKPND 69
 | : | : || |
 Db 2414 LPTTTQKPTD 2423

RESULT 9

HIS7_PHYPR

ID HIS7_PHYPR STANDARD; PRT; 452 AA.
 AC P28624;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19) (IGPD).
 GN HIS3.
 OS Phytophthora parasitica (Potato buckeye rot agent).
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 OC Phytophthora.
 OX NCBI_TaxID=4792;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1829;
 RA Baltrusch-Weiter M., Karlovsky P., Prell H.H.;
 RL Submitted (JAN-1992) to the EMBL/GenBank/DDBJ databases.
 CC -!- CATALYTIC ACTIVITY: D-erythro-1-(imidazol-4-yl)glycerol 3-
 CC phosphate = 3-(imidazol-4-yl)-2-oxopropyl phosphate + H(2)O.
 CC -!- PATHWAY: Histidine biosynthesis; sixth step.
 CC -!- SIMILARITY: Belongs to the imidazoleglycerol-phosphate dehydratase
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Z11591; CAA77675.1; -.
 DR PIR; S22199; S22199.
 DR InterPro; IPR006438; HAD-SF-IA-hyp1.
 DR InterPro; IPR006543; Histidinol-phos.
 DR InterPro; IPR005834; Hydrolase.
 DR InterPro; IPR000807; IGPD.
 DR Pfam; PF00702; Hydrolase; 1.
 DR Pfam; PF00475; IGPD; 1.
 DR ProDom; PD002282; IGPD; 1.
 DR TIGRFAMs; TIGR01548; HAD-SF-IA-hyp1; 1.
 DR TIGRFAMs; TIGR01656; Histidinol-ppas; 1.
 DR PROSITE; PS00954; IGP_DEHYDRATASE_1; 1.
 DR PROSITE; PS00955; IGP_DEHYDRATASE_2; 1.
 KW Histidine biosynthesis; Lyase; Multifunctional enzyme.
 FT DOMAIN 1 233 UNKNOWN ACTIVITY.
 FT DOMAIN 234 452 IMIDAZOLEGLYCEROL-PHOSPHATE DEHYDRATASE.
 SQ SEQUENCE 452 AA; 47961 MW; CAE66BE32A9E53A1 CRC64;

Query Match 19.7%; Score 74; DB 1; Length 452;
 Best Local Similarity 34.4%; Pred. No. 5.3;
 Matches 21; Conservative 8; Mismatches 22; Indels 10; Gaps 3;

Qy 12 QVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKP-----DVPPKPSFAPLSTSM 65
 Db 111 ELHRRQPKGMAVVTGR-PRKDCAKFLTTHGIE---DLFPVQIWLEDCPPKPSPEPILLAL 166
 Qy 66 K 66
 Db 167 K 167

RESULT 10

M4K3_HUMAN

ID M4K3_HUMAN STANDARD; PRT; 894 AA.
 AC Q8IVH8; Q8IVH7; Q9UDM5; Q9Y6R5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 3 (EC 2.7.1.37)
 DE (MAPK/ERK kinase kinase kinase 3) (MEK kinase kinase 3) (MEKKK 3)
 DE (Germinal center kinase related protein kinase) (GLK).
 GN MAP4K3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, TISSUE SPECIFICITY, AND
 RP MUTAGENESIS OF LYS-48.
 RC TISSUE=Macrophage, and Skeletal muscle;
 RX MEDLINE=97420743; PubMed=9275185;
 RA Diener K., Wang X.S., Chen C., Meyer C.F., Keesler G., Zukowski M.,
 RA Tan T.-H., Yao Z.;
 RT "Activation of the c-Jun N-terminal kinase pathway by a novel protein
 RT kinase related to human germinal center kinase."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:9687-9692(1997).
 RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RA Gorry M.C., Zhang Y., Marks J.J., Suppe B., Hart S., Cortelli J.,
 RA Pallos D., Hart T.C.;
 RT "Physical/genetic map of the 2p22-2p21 region on chromosome 2.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE OF 1-712 FROM N.A. (ISOFORM 1).
 RA Edwards J., Wohldmann P., Hawkins M., Harkins R.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: May play a role in the response to environmental stress.
 CC Appears to act upstream of the c-jun N-terminal
 CC pathway.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a
 CC phosphoprotein.
 CC -!- COFACTOR: Magnesium.
 CC -!- SUBUNIT: Interacts with SH3GL2. Interaction appears to regulate
 CC MAP4K3-mediated JNK activation (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q8IVH8-1; Sequence=Displayed;
 CC Note=No experimental confirmation available;
 CC Name=2;
 CC IsoId=Q8IVH8-2; Sequence=VSP_007052;
 CC Name=3;
 CC IsoId=Q8IVH8-3; Sequence=VSP_007053;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed in all tissues.
 CC examined, with high levels in heart, brain, placenta, skeletal
 CC muscle, kidney and pancreas and lower levels in lung and
 CC liver.
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC STE20 subfamily.
 CC -!- SIMILARITY: Contains 1 CNH domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF000145; AAC15472.1; -.
 DR EMBL; AF445413; AAN75849.1; -.
 DR EMBL; AF445385; AAN75849.1; JOINED.
 DR EMBL; AF445386; AAN75849.1; JOINED.
 DR EMBL; AF445387; AAN75849.1; JOINED.
 DR EMBL; AF445388; AAN75849.1; JOINED.
 DR EMBL; AF445390; AAN75849.1; JOINED.
 DR EMBL; AF445391; AAN75849.1; JOINED.
 DR EMBL; AF445392; AAN75849.1; JOINED.
 DR EMBL; AF445393; AAN75849.1; JOINED.
 DR EMBL; AF445394; AAN75849.1; JOINED.
 DR EMBL; AF445395; AAN75849.1; JOINED.
 DR EMBL; AF445396; AAN75849.1; JOINED.
 DR EMBL; AF445397; AAN75849.1; JOINED.

DR EMBL; AF445398; AAN75849.1; JOINED.
 DR EMBL; AF445399; AAN75849.1; JOINED.
 DR EMBL; AF445400; AAN75849.1; JOINED.
 DR EMBL; AF445401; AAN75849.1; JOINED.
 DR EMBL; AF445402; AAN75849.1; JOINED.
 DR EMBL; AF445403; AAN75849.1; JOINED.
 DR EMBL; AF445404; AAN75849.1; JOINED.
 DR EMBL; AF445405; AAN75849.1; JOINED.
 DR EMBL; AF445406; AAN75849.1; JOINED.
 DR EMBL; AF445407; AAN75849.1; JOINED.
 DR EMBL; AF445408; AAN75849.1; JOINED.
 DR EMBL; AF445409; AAN75849.1; JOINED.
 DR EMBL; AF445410; AAN75849.1; JOINED.
 DR EMBL; AF445411; AAN75849.1; JOINED.
 DR EMBL; AF445412; AAN75849.1; JOINED.
 DR EMBL; AF445413; AAN75850.1; -.
 DR EMBL; AF445385; AAN75850.1; JOINED.
 DR EMBL; AF445386; AAN75850.1; JOINED.
 DR EMBL; AF445387; AAN75850.1; JOINED.
 DR EMBL; AF445388; AAN75850.1; JOINED.
 DR EMBL; AF445390; AAN75850.1; JOINED.
 DR EMBL; AF445391; AAN75850.1; JOINED.
 DR EMBL; AF445392; AAN75850.1; JOINED.
 DR EMBL; AF445393; AAN75850.1; JOINED.
 DR EMBL; AF445394; AAN75850.1; JOINED.
 DR EMBL; AF445395; AAN75850.1; JOINED.
 DR EMBL; AF445397; AAN75850.1; JOINED.
 DR EMBL; AF445398; AAN75850.1; JOINED.
 DR EMBL; AF445399; AAN75850.1; JOINED.
 DR EMBL; AF445400; AAN75850.1; JOINED.
 DR EMBL; AF445401; AAN75850.1; JOINED.
 DR EMBL; AF445402; AAN75850.1; JOINED.
 DR EMBL; AF445403; AAN75850.1; JOINED.
 DR EMBL; AF445404; AAN75850.1; JOINED.
 DR EMBL; AF445405; AAN75850.1; JOINED.
 DR EMBL; AF445406; AAN75850.1; JOINED.
 DR EMBL; AF445407; AAN75850.1; JOINED.
 DR EMBL; AF445408; AAN75850.1; JOINED.
 DR EMBL; AF445409; AAN75850.1; JOINED.
 DR EMBL; AF445410; AAN75850.1; JOINED.
 DR EMBL; AF445411; AAN75850.1; JOINED.
 DR EMBL; AF445412; AAN75850.1; JOINED.
 DR EMBL; AC007684; AAF19240.1; -.
 DR HSSP; P24941; 1B38.
 DR Genew; HGNC:6865; MAP4K3.
 DR MIM; 604921; -.
 DR GO; GO:0005524; F:ATP binding; IDA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
 DR GO; GO:0007243; P:protein kinase cascade; IDA.
 DR GO; GO:0006950; P:response to stress; IDA.
 DR InterPro; IPR001180; Citron.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD0000001; Prot_kinase; 1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 KW ATP-binding; Transferase; Serine/threonine-protein kinase;
 KW Alternative splicing.
 FT DOMAIN 16 273 PROTEIN KINASE.
 FT DOMAIN 562 874 CNH.
 FT NP_BIND 22 30 ATP (BY SIMILARITY).
 FT BINDING 48 48 ATP.
 FT ACT_SITE 136 136 BY SIMILARITY.
 FT VARSPLIC 1 12 MNPGFDLSRRNP -> MA (in isoform 2).
 FT /FTid=VSP_007052.
 FT VARSPLIC 352 372 Missing (in isoform 3).
 FT /FTid=VSP_007053.
 FT MUTAGEN 48 48 K->E: LOSS OF KINASE ACTIVITY AND ABILITY
 FT TO ACTIVATE JNK FAMILY.
 FT CONFLICT 392 392 N -> D (IN REF. 1; AAC15472).
 SQ SEQUENCE 894 AA; 101315 MW; 6EB77BBB34E5B733 CRC64;

Query Match 19.7%; Score 74; DB 1; Length 894;
 Best Local Similarity 32.8%; Pred. No. 11;
 Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;

Qy 1 PPPAPQRVDSI----QVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKP 56
 ||| | : || :||:: | | : | | || || | || |
 Db 432 PPPLPPKPKSIFIPQEMHSTEDENQG-TIKRCP-----MSGSPAKPSQVPPRPPPP 481
 Qy 57 SFAP 60
 |
 Db 482 RLPP 485

RESULT 11

MINT_HUMAN

ID MINT_HUMAN STANDARD; PRT; 3664 AA.
 AC Q96T58; Q9H9A8; Q9NWH5; Q9UQ01; Q9Y556;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Msx2-interacting protein (SMART/HDAC1 associated repressor protein).
 GN MINT OR SHARP OR KIAA0929.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, INDUCTION,
 RP RNA-BINDING, AND INTERACTION WITH NCOR2; HDAC1; HDAC2; RBBP4; MBD3;
 RP RAR AND MTAL1.
 RC TISSUE=Liver, and Pituitary;
 RX MEDLINE=21231190; PubMed=11331609;
 RA Shi Y., Downes M., Xie W., Kao H.-Y., Ordentlich P., Tsai C.-C.,
 RA Hon M., Evans R.M.;
 RT "Sharp, an inducible cofactor that integrates nuclear receptor

RT repression and activation.";
 RL Genes Dev. 15:1140-1151(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bird C.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE OF 294-3664 FROM N.A.
 RA Rhodes S., Huckle E.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE OF 793-1595 FROM N.A., AND VARIANT PRO-1091.
 RC TISSUE=Embryo, and Teratocarcinoma;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
 RN [5]
 RP SEQUENCE OF 2002-3664 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99246063; PubMed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 RN [6]
 RP INTERACTION WITH PPARD.
 RX MEDLINE=21874127; PubMed=11867749;
 RA Shi Y., Hon M., Evans R.M.;
 RT "The peroxisome proliferator-activated receptor delta, an integrator
 RT of transcriptional repression and nuclear receptor signaling.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:2613-2618(2002).
 RN [7]
 RP FUNCTION, AND INTERACTION WITH RBPSUH.
 RX MEDLINE=22261914; PubMed=12374742;
 RA Oswald F., Kostezka U., Astrahantseff K., Bourteele S., Dillinger K.,
 RA Zechner U., Ludwig L., Wilda M., Hameister H., Knoechel W., Liptay S.,
 RA Schmid R.M.;
 RT "SHARP is a novel component of the Notch/RBP-Jkappa signalling
 RT pathway.";
 RL EMBO J. 21:5417-5426(2002).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF SPOC DOMAIN.
 RX MEDLINE=22777836; PubMed=12897056;
 RA Ariyoshi M., Schwabe J.W.R.;
 RT "A conserved structural motif reveals the essential transcriptional
 RT repression function of Spen proteins and their role in developmental
 RT signaling.";
 RL Genes Dev. 17:1909-1920(2003).
 CC -!- FUNCTION: Essential corepressor protein, which probably regulates
 CC different key pathways such as the Notch pathway. Negative
 CC regulator of the Notch pathway via its interaction with RBPSUH,

CC which prevents the association between NOTCH1 and RBPSUH, and
 CC therefore suppresses the transactivation activity of Notch
 CC signaling. Blocks the differentiation of precursor B cells into
 CC marginal zone B cells. Probably represses transcription via the
 CC recruitment of large complexes containing histone deacetylase
 CC proteins. May bind both to DNA and RNA.
 CC -!- SUBUNIT: Interacts with MSX2 (By similarity). Interacts with
 CC NCOR2, HDAC1, HDAC2, RBBP4, MBD3 and MTA1L1. Interacts with
 CC RBPSUH; this interaction may prevent the interaction between
 CC RBPSUH and NOTCH1. Interacts with the nuclear receptors RAR and
 CC PPARG. Interacts with RAR in absence of ligand. Bind to the
 CC steroid receptor RNA coactivator SRA.
 CC -!- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
 CC -!- TISSUE SPECIFICITY: Expressed at high level in brain, testis,
 CC spleen and thymus. Expressed at intermediate level in kidney,
 CC liver, mammary gland and skin.
 CC -!- INDUCTION: By hormone 17-beta-estradiol (E2).
 CC -!- DOMAIN: The RID domain mediates the interaction with nuclear
 CC receptors (By similarity).
 CC -!- DOMAIN: The SPOC domain, which mediates the interaction with
 CC NCOR2, is essential for the repressive activity.
 CC -!- SIMILARITY: Belongs to the Spen family.
 CC -!- SIMILARITY: Contains 1 RID (receptor interacting) domain.
 CC -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
 CC -!- SIMILARITY: Contains 1 SPOC domain.
 CC -!- CAUTION: Ref.2 sequences differ from that shown due to erroneous
 CC gene model prediction.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF356524; AAK52750.1; -.
 DR EMBL; AL034555; CAB85442.1; ALT_SEQ.
 DR EMBL; AL034555; CAB85444.1; ALT_SEQ.
 DR EMBL; AL450998; -; NOT_ANNOTATED_CDS.
 DR EMBL; AL096858; CAB51072.1; ALT_INIT.
 DR EMBL; AK000882; BAA91405.1; ALT_INIT.
 DR EMBL; AK022949; BAB14324.1; ALT_INIT.
 DR EMBL; AB023146; BAA76773.1; -.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR PDB; 1OW1; 19-AUG-03.
 DR Pfam; PF00076; rrm; 4.
 DR SMART; SM00360; RRM; 4.
 DR PROSITE; PS50102; RRM; 4.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 DR PROSITE; PS50917; SPOC; 1.
 KW Transcription regulation; Repressor; Nuclear protein; DNA-binding;
 KW Repeat; RNA-binding; Coiled coil; 3D-structure; Polymorphism.
 FT DOMAIN 1 573 DNA-BINDING (BY SIMILARITY).
 FT DOMAIN 6 81 RNA-BINDING (RRM) 1.
 FT DOMAIN 335 415 RNA-BINDING (RRM) 2.
 FT DOMAIN 438 513 RNA-BINDING (RRM) 3.

FT	DOMAIN	517	589	RNA-BINDING (RRM) 4.
FT	DOMAIN	688	715	COILED COIL (POTENTIAL).
FT	DOMAIN	977	1004	COILED COIL (POTENTIAL).
FT	DOMAIN	1170	1191	COILED COIL (POTENTIAL).
FT	DOMAIN	1408	1428	COILED COIL (POTENTIAL).
FT	DOMAIN	1496	1529	COILED COIL (POTENTIAL).
FT	DOMAIN	1592	1612	COILED COIL (POTENTIAL).
FT	DOMAIN	1928	1944	COILED COIL (POTENTIAL).
FT	DOMAIN	2201	2707	RID.
FT	DOMAIN	3498	3664	SPOC.
FT	DOMAIN	2130	2464	INTERACTION WITH MSX2 (BY SIMILARITY).
FT	DOMAIN	2709	2870	INTERACTION WITH RBPSUH (BY SIMILARITY).
FT	DOMAIN	125	277	ARG-RICH.
FT	DOMAIN	240	325	SER-RICH.
FT	DOMAIN	616	810	ARG-RICH.
FT	DOMAIN	624	697	TYR-RICH.
FT	DOMAIN	2428	2520	PRO-RICH.
FT	DOMAIN	3220	3482	PRO-RICH.
FT	VARIANT	970	970	A -> V (in dbSNP:848208).
FT				/FTId=VAR_017119.
FT	VARIANT	1091	1091	L -> P (in dbSNP:848209).
FT				/FTId=VAR_017120.
FT	VARIANT	2360	2360	N -> D (in dbSNP:848210).
FT				/FTId=VAR_017121.
FT	CONFLICT	956	956	G -> D (IN REF. 4).
SQ	SEQUENCE	3664	AA; 402245	MW; 5228C58533E5B27B CRC64;

Query Match 19.5%; Score 73.5; DB 1; Length 3664;
 Best Local Similarity 37.5%; Pred. No. 62;
 Matches 24; Conservative 9; Mismatches 26; Indels 5; Gaps 3;

Qy	2	PPAPQ-RVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTP---SLKPDVPPPKPS	57
		: : : : : : : :	
Db	2378	PEAPQEEKQSEKPHSTPPQSCSDLSKIPS-TENSSQEISVEERTPTKASVPPDLPPPPQ	2436

Qy	58	FAPL	61
		:	
Db	2437	PAPV	2440

RESULT 12

PCLO_RAT

ID	PCLO_RAT	STANDARD;	PRT; 5085	AA.
AC	Q9JKS6; Q9JLT1;			
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	Piccolo protein	(Multidomain presynaptic cytomatrix protein).		
GN	PCLO.			
OS	Rattus norvegicus	(Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH RABAC1.			
RX	MEDLINE=20170257; PubMed=10707984;			
RA	Fenster S.D., Chung W.J., Zhai R., Cases-Langhoff C., Voss B.,			

RA Garner A.M., Kaempf U., Kindler S., Gundelfinger E.D., Garner C.C.;
 RT "Piccolo, a presynaptic zinc finger protein structurally related to
 RT bassoon.";
 RL Neuron 25:203-214(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Fenster S.D., Cases-Langhoff C., Gundelfinger E.D., Garner C.C.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP CALICUM-BINDING ACTIVITY, AND MUTAGENESIS OF ASP-4668; ASP-4674;
 RP VAL-4688; MET-4689; VAL-4690; SER-4691; GLN-4692; ASN-4693 AND
 RP ALA-4694.
 RX MEDLINE=21181819; PubMed=11285225;
 RA Gerber S.H., Garcia J., Rizo J., Suedhof T.C.;
 RT "An unusual C(2)-domain in the active-zone protein piccolo:
 RT implications for Ca(2+) regulation of neurotransmitter release.";
 RL EMBO J. 20:1605-1619(2001).
 CC -!- FUNCTION: May act as a scaffolding protein involved in the
 CC organization of synaptic active zones and in synaptic vesicle
 CC trafficking (By similarity).
 CC -!- SUBUNIT: Interacts with Rabac1/Pral, RIMS2 and profilin (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Concentrated at presynaptic side of synaptic
 CC junctions.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9JKS6-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9JKS6-2; Sequence=VSP_003930, VSP_003931;
 CC -!- DOMAIN: C2 domain 1 is involved in binding calcium and
 CC phospholipids. Calcium binds with low affinity but with high
 CC specificity and induces a large conformational change.
 CC -!- SIMILARITY: Contains 2 C2 domains.
 CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF138789; AAF07822.2; -.
 DR EMBL; AF227534; AAF63196.1; -.
 DR HSSP; P04410; 1A25.
 DR GO; GO:0045202; C:synaptic junction; IDA.
 DR GO; GO:0005509; F:calcium ion binding; IDA.
 DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IDA.
 DR GO; GO:0005522; F:profilin binding; ISS.
 DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
 DR GO; GO:0016080; P:synaptic vesicle targeting; NAS.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR008899; Znf_piccolo.
 DR Pfam; PF00168; C2; 2.

DR Pfam; PF00595; PDZ; 1.
 DR Pfam; PF05715; Zf_piccolo; 2.
 DR SMART; SM00239; C2; 2.
 DR SMART; SM00228; PDZ; 1.
 DR PROSITE; PS00499; C2_DOMAIN_1; 1.
 DR PROSITE; PS50004; C2_DOMAIN_2; 2.
 DR PROSITE; PS50106; PDZ; 1.
 KW Calcium/phospholipid-binding; Metal-binding; Zinc; Zinc-finger;
 KW Repeat; Alternative splicing.
 FT DOMAIN 372 491 12 X 10 AA TANDEM APPROXIMATE REPEATS OF
 FT P-A-K-P-Q-P-Q-Q-P-X.
 FT ZN_FING 523 547 C4-TYPE (POTENTIAL).
 FT ZN_FING 1010 1033 C4-TYPE (POTENTIAL).
 FT DOMAIN 2351 2362 POLY-PRO.
 FT DOMAIN 4442 4536 PDZ.
 FT DOMAIN 4653 4752 C2 DOMAIN 1.
 FT DOMAIN 4968 5059 C2 DOMAIN 2.
 FT VARSPLIC 4876 4880 TKPTN -> SKRRK (in isoform 2).
 FT /FTid=VSP_003930.
 FT VARSPLIC 4881 5085 Missing (in isoform 2).
 FT /FTid=VSP_003931.
 FT MUTAGEN 4668 4668 D->A: COMPLETE LOSS OF CALCIUM-BINDING
 FT AND CALCIUM-DEPENDENT PHOSPHOLIPID
 FT BINDING ACTIVITY.
 FT MUTAGEN 4674 4674 D->A: COMPLETE LOSS OF CALCIUM-BINDING
 FT AND CALCIUM-DEPENDENT PHOSPHOLIPID
 FT BINDING ACTIVITY.
 FT MUTAGEN 4688 4688 V->S: SMALL INCREASE IN AFFINITY FOR
 FT CALCIUM.
 FT MUTAGEN 4688 4689 VM->SS: 10-FOLD INCREASE IN AFFINITY FOR
 FT CALCIUM.
 FT MUTAGEN 4689 4689 M->S: INCREASED AFFINITY FOR CALCIUM.
 FT MUTAGEN 4690 4691 VV->SS: 10-FOLD INCREASE IN AFFINITY FOR
 FT CALCIUM.
 FT MUTAGEN 4692 4693 QN->AA: MODERATE INCREASE IN AFFINITY FOR
 FT CALCIUM.
 FT MUTAGEN 4694 4694 A->S: NO EFFECT ON CALCIUM-BINDING
 FT ACTIVITY.
 SQ SEQUENCE 5085 AA; 552702 MW; 5A1BB543201A7450 CRC64;

Query Match 19.5%; Score 73.5; DB 1; Length 5085;
 Best Local Similarity 26.4%; Pred. No. 90;
 Matches 19; Conservative 8; Mismatches 20; Indels 25; Gaps 3;

Qy 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDPVPPKPSFAP 60
 ||| | : | | : | | | : || : ||| :
 Db 2432 PPPVPPKPSQI-----PTGLVFT-----HRPEAIKPPPIAPKPAVPQ 2467
 Qy 61 LS-TSMKPNDAC 71
 : | : || | |
 Db 2468 IPVTTQKPTDTC 2479

RESULT 13

MLL3_HUMAN

ID MLL3_HUMAN STANDARD; PRT; 4911 AA.
 AC Q8NEZ4; Q8NC02; Q8NDF6; Q9H9P4; Q9NR13; Q9P222; Q9UDR7;

DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Myeloid/lymphoid or mixed-lineage leukemia protein 3 homolog (Histone-
 DE lysine N-methyltransferase, H3 lysine-4 specific MLL3) (EC 2.1.1.43)
 DE (Homologous to ALR protein).
 GN MLL3 OR HALR OR KIAA1506.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal thymus;
 RX MEDLINE=21888622; PubMed=11891048;
 RA Ruault M., Brun M.-E., Ventura M., Roizes G., De Sario A.;
 RT "MLL3, a new human member of the TRX/MLL gene family, maps to 7q36, a
 RT chromosome region frequently deleted in myeloid leukaemia."
 RL Gene 284:73-81(2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Cervical carcinoma;
 RX MEDLINE=21574953; PubMed=11718452;
 RA Tan Y.C., Chow V.T.;
 RT "Novel human HALR (MLL3) gene encodes a protein homologous to ALR and
 RT to ALL-1 involved in leukemia, and maps to chromosome 7q36 associated
 RT with leukemia and developmental defects."
 RL Cancer Detect. Prev. 25:454-469(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22737999; PubMed=12853948;
 RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
 RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
 RA Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
 RA Fewell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
 RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
 RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
 RA Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
 RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
 RA Kozlowski-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
 RA Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strowmatt C.,
 RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
 RA Wendl M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,
 RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
 RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
 RA Clifton S.W., Chissole S.L., Marra M.A., Raymond C., Haugen E.,
 RA Gillett W., Zhou Y., James R., Phelps K., Iadanoto S., Bubb K.,
 RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
 RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
 RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
 RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
 RA Waterston R.H., Wilson R.K.;
 RT "The DNA sequence of human chromosome 7."
 RL Nature 424:157-164(2003).
 RN [4]
 RP SEQUENCE OF 556-3865 FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;

RX MEDLINE=20277482; PubMed=10819331;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XVII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:143-150(2000).
 RN [5]
 RP SEQUENCE OF 3193-3865 AND 4460-4911 FROM N.A.
 RC TISSUE=Placenta;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
 RN [6]
 RP SEQUENCE OF 3879-4911 FROM N.A.
 RC TISSUE=Testis;
 RA Duesterhoeft A., Lauber J., Mewes H.-W., Weil B., Wiemann S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
 RN [7]
 RP INTERACTION WITH ASC-2/NCOA6 CONTAINING COMPLEX (ISOFORM 2).
 RC TISSUE=Cervical carcinoma;
 RX MEDLINE=22371496; PubMed=12482968;
 RA Goo Y.-H., Sohn Y.C., Kim D.-H., Kim S.-W., Kang M.-J., Jung D.-J.,
 RA Kwak E., Barlev N.A., Berger S.L., Chow V.T., Roeder R.G.,
 RA Azorsa D.O., Meltzer P.S., Suh P.-G., Song E.J., Lee K.-J., Lee Y.C.,
 RA Lee J.W.;
 RT "Activating signal cointegrator 2 belongs to a novel steady-state
 RT complex that contains a subset of trithorax group proteins.";
 RL Mol. Cell. Biol. 23:140-149(2003).
 CC -!- FUNCTION: Belongs to the ASC-2/NCOA6 complex (ASCOM), a
 CC coactivator complex of nuclear receptors, involved in
 CC transcriptional coactivation. MLL3 may be a catalytic subunit of
 CC this complex, which weakly methylates Lys-4 of histone H3. This is
 CC a specific tag for epigenetic transcriptional activation. May be
 CC involved in leukemogenesis and developmental disorder.
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
 CC S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
 CC -!- SUBUNIT: Belongs to the ASC-2/NCOA6 complex (ASCOM), which
 CC contains ASC-2/NCOA6, the retinoblastoma-binding protein RBQ-3/
 CC RBBP5, alpha- and beta-tubulins, the trithorax group proteins
 CC MLL2 and MLL3, and ASH2/ASCL2. Interacts with histone H3.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8NEZ4-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8NEZ4-2; Sequence=VSP_008561, VSP_008562;
 CC -!- TISSUE SPECIFICITY: Highly expressed in testis and ovary, followed
 CC by brain and liver. Also expressed in placenta, peripheral
 CC blood, fetal thymus, heart, lung and kidney. Within brain,
 CC expression was highest in hippocampus, caudate nucleus, and

CC substantia nigra. Not detected in skeletal muscle and fetal liver.
 CC -!- DOMAIN: The SET domain interacts with histone H3 but not H2A, H2B
 CC and H4, and may have a H3 lysine specific methylation activity.
 CC -!- MISCELLANEOUS: Found in a critical region of chromosome 7, which
 CC is commonly deleted in malignant myeloid disorders. Partial
 CC duplication of the MLL3 gene are found in the juxtacentromeric
 CC region of chromosomes 1, 2, 13 and 21. Juxtacentromeric
 CC reshuffling of the MLL3 gene has generated the BAGE genes.
 CC -!- SIMILARITY: Belongs to the TRX/MLL family.
 CC -!- SIMILARITY: Contains 1 DHHC-type zinc finger.
 CC -!- SIMILARITY: Contains 6 PHD-type zinc fingers.
 CC -!- SIMILARITY: Contains 1 post-SET domain.
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -!- SIMILARITY: Contains 1 SET domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AY024361; AAK00583.1; -.
 DR EMBL; AF264750; AAF74766.2; -.
 DR EMBL; AC006017; AAD45822.1; -.
 DR EMBL; AC104692; -; NOT_ANNOTATED_CDS.
 DR EMBL; AC005631; -; NOT_ANNOTATED_CDS.
 DR EMBL; AB040939; BAA96030.2; -.
 DR EMBL; AK022687; BAB14179.1; -.
 DR EMBL; AK075113; BAC11409.1; -.
 DR EMBL; AL833924; CAD38780.1; -.
 DR Genew; HGNC:13726; MLL3.
 DR MIM; 606833; -.
 DR InterPro; IPR000637; AT_hook.
 DR InterPro; IPR003889; FYrich_C.
 DR InterPro; IPR003888; FYrich_N.
 DR InterPro; IPR000910; HMG_12_box.
 DR InterPro; IPR003616; PostSET.
 DR InterPro; IPR001214; SET.
 DR InterPro; IPR001594; Znf_DHHC.
 DR InterPro; IPR001965; Znf_PHD.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00505; HMG_box; 1.
 DR Pfam; PF00628; PHD; 6.
 DR Pfam; PF00856; SET; 1.
 DR SMART; SM00542; FYRC; 1.
 DR SMART; SM00541; FYRN; 1.
 DR SMART; SM00398; HMG; 1.
 DR SMART; SM00249; PHD; 8.
 DR SMART; SM00508; PostSET; 1.
 DR SMART; SM00317; SET; 1.
 DR PROSITE; PS00354; HMGI_Y; 1.
 DR PROSITE; PS50868; POST_SET; 1.
 DR PROSITE; PS50280; SET; 1.
 DR PROSITE; PS50216; ZF_DHHC; 1.
 DR PROSITE; PS01359; ZF_PHD_1; 5.

DR PROSITE; PS50016; ZF_PHD_2; 6.
 DR PROSITE; PS50089; ZF_RING_2; 1.
 KW Transferase; Methyltransferase; Chromatin regulator; Activator;
 KW DNA-binding; Nuclear protein; Transcription regulation; Coiled coil;
 KW Zinc-finger; Repeat; Alternative splicing; Polymorphism.

FT	ZN_FING	341	391	PHD-TYPE 1.
FT	ZN_FING	344	389	RING-TYPE.
FT	ZN_FING	388	438	PHD-TYPE 2.
FT	ZN_FING	436	489	DHHC-TYPE.
FT	ZN_FING	464	520	PHD-TYPE 3.
FT	ZN_FING	957	1010	PHD-TYPE 4.
FT	ZN_FING	1007	1057	PHD-TYPE 5.
FT	ZN_FING	1084	1139	PHD-TYPE 6.
FT	DOMAIN	4770	4891	SET.
FT	DOMAIN	4895	4911	POST-SET.
FT	DOMAIN	92	112	COILED COIL (POTENTIAL).
FT	DOMAIN	644	672	COILED COIL (POTENTIAL).
FT	DOMAIN	1338	1366	COILED COIL (POTENTIAL).
FT	DOMAIN	1754	1787	COILED COIL (POTENTIAL).
FT	DOMAIN	3054	3081	COILED COIL (POTENTIAL).
FT	DOMAIN	3173	3272	COILED COIL (POTENTIAL).
FT	DOMAIN	3391	3433	COILED COIL (POTENTIAL).
FT	DNA_BIND	34	46	A.T HOOK (BY SIMILARITY).
FT	DOMAIN	1719	1796	GLN-RICH.
FT	DOMAIN	1834	2281	PRO-RICH.
FT	DOMAIN	2412	2630	PRO-RICH.
FT	DOMAIN	2690	2786	ASP-RICH.

Query Match 19.4%; Score 73; DB 1; Length 4911;
 Best Local Similarity 32.5%; Pred. No. 97;
 Matches 27; Conservative 11; Mismatches 27; Indels 18; Gaps 6;

Qy 1 PPPAPQRV---DSI-QVHSSQPSGQAV-----TVSRQPS-LNAYNSLTRSGLK RTP---- 46
 ||||| |: ||: | :|| | : || || :: | : | ||
 Db 1855 PPPAPSRIPIQDSL SQAQTSQPPSPQVFSPGSSNSRPPSPMDPYAKMV--GTPRPPPVGH 1912

Qy 47 --SLKPDVPPKPSFAPLSTSMKP 67
 | : | : |||: :|
 Db 1913 SFSRRNSAAPVENCTPLSSVSRP 1935

RESULT 14

V70K_TYMV

ID V70K TYMV STANDARD; PRT; 628 AA.

AC P10357;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 01-AUG-1992 (Rel. 23, Last annotation update)

DE 69 kDa protein.

OS Turnip yellow mosaic virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymoviridae;

OC Tymovirus.

OX NCBI_TaxID=12154;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88289359; PubMed=3399388;

RA Morch M.D., Boyer J.C., Haenni A.L.;

RT "Overlapping open reading frames revealed by complete nucleotide
RT sequencing of turnip yellow mosaic virus genomic RNA.";
RL Nucleic Acids Res. 16:6157-6173(1988).
CC -!- FUNCTION: Not known.
CC -!- SIMILARITY: TO 65 TO 70 kDa PROTEIN FROM OTHER TYMOVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07441; CAA30321.1; ALT_SEQ.
DR PIR; S01955; S01955.
DR InterPro; IPR004935; Tymo_45_70kDa.
DR Pfam; PF03251; Tymo_45kd_70kd; 1.
SQ SEQUENCE 628 AA; 69195 MW; 9B01CE5ADFCEAC77 CRC64;

Query Match 19.3%; Score 72.5; DB 1; Length 628;
Best Local Similarity 29.6%; Pred. No. 11;
Matches 21; Conservative 7; Mismatches 14; Indels 29; Gaps 3;

Qy 2 PPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTP----SLKPDV-PPKP 56
| | | | | | : : | : : | | | : : | | | | |
Db 119 PPAPQRQHSPLPLHITRPS-----RFPHHFHARRPDVLPSPV 154

Qy 57 SFAPLSTSMKP 67
| : | | |
Db 155 DHGPVLTETKP 165

RESULT 15

MLL2_HUMAN

ID MLL2_HUMAN STANDARD; PRT; 5262 AA.
AC O14686; O14687;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myeloid/lymphoid or mixed-lineage leukemia protein 2 (ALL1-related
DE protein).
GN MLL2 OR ALR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX MEDLINE=97388474; PubMed=9247308;
RA Prasad R., Zhadanov A.B., Sedkov Y., Bullrich F., Druck T.,
RA Rallapalli R., Yano T., Alder H., Croce C.M., Huebner K., Mazo A.,
RA Canaani E.;
RT "Structure and expression pattern of human ALR, a novel gene with
RT strong homology to ALL-1 involved in acute leukemia and to Drosophila
RT trithorax.";
RL Oncogene 15:549-560(1997).

RN [2]
 RP INTERACTION WITH ASC-2/NCOA6 CONTAINING COMPLEX.
 RC TISSUE=Cervical carcinoma;
 RX MEDLINE=22371496; PubMed=12482968;
 RA Goo Y.-H., Sohn Y.C., Kim D.-H., Kim S.-W., Kang M.-J., Jung D.-J.,
 RA Kwak E., Barlev N.A., Berger S.L., Chow V.T., Roeder R.G.,
 RA Azorsa D.O., Meltzer P.S., Suh P.-G., Song E.J., Lee K.-J., Lee Y.C.,
 RA Lee J.W.;
 RT "Activating signal cointegrator 2 belongs to a novel steady-state
 RT complex that contains a subset of trithorax group proteins."
 RL Mol. Cell. Biol. 23:140-149(2003).
 CC -!- FUNCTION: May be involved in transcriptional regulation.
 CC -!- SUBUNIT: Belongs to the ASC-2/NCOA6 complex (ASCOM), which
 CC contains ASC-2/NCOA6, the retinoblastoma-binding protein RBQ-3/
 CC RBBP5, alpha- and beta-tubulins, the trithorax group proteins
 CC MLL2 and MLL3, and ASH2/ASCL2.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=O14686-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O14686-2; Sequence=VSP_008563, VSP_008559;
 CC Name=3;
 CC IsoId=O14686-3; Sequence=VSP_008560;
 CC -!- TISSUE SPECIFICITY: Expressed in most adult tissues, including a
 CC variety of hematoipoietic cells, with the exception of the liver.
 CC -!- MISCELLANEOUS: This gene mapped to a chromosomal region involved
 CC in duplications and translocations associated with cancer.
 CC -!- SIMILARITY: Belongs to the transcription factor trithorax family.
 CC -!- SIMILARITY: Contains 5 PHD-type zinc fingers.
 CC -!- SIMILARITY: Contains 1 post-SET domain.
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -!- SIMILARITY: Contains 1 SET domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF010403; AAC51734.1; -.
 DR EMBL; AF010404; AAC51735.1; -.
 DR PIR; T03454; T03454.
 DR PIR; T03455; T03455.
 DR Genew; HGNC:7133; MLL2.
 DR MIM; 602113; -.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0003700; F:transcription factor activity; TAS.
 DR GO; GO:0007048; P:oncogenesis; TAS.
 DR GO; GO:0006366; P:transcription from Pol II promoter; TAS.
 DR InterPro; IPR003889; FYrich_C.
 DR InterPro; IPR003888; FYrich_N.
 DR InterPro; IPR000910; HMG_12_box.
 DR InterPro; IPR003616; PostSET.

DR InterPro; IPR006118; Recombinase.
 DR InterPro; IPR001214; SET.
 DR InterPro; IPR001965; Znf_PHD.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00628; PHD; 5.
 DR Pfam; PF00856; SET; 1.
 DR SMART; SM00542; FYRC; 1.
 DR SMART; SM00541; FYRN; 1.
 DR SMART; SM00398; HMG; 1.
 DR SMART; SM00249; PHD; 7.
 DR SMART; SM00508; PostSET; 1.
 DR SMART; SM00184; RING; 3.
 DR SMART; SM00317; SET; 1.
 DR PROSITE; PS50868; POST_SET; 1.
 DR PROSITE; PS50280; SET; 1.
 DR PROSITE; PS01359; ZF_PHD_1; 5.
 DR PROSITE; PS50016; ZF_PHD_2; 5.
 DR PROSITE; PS50089; ZF_RING_2; 1.
 KW Nuclear protein; Transcription regulation; Coiled coil; Zinc-finger;
 KW Repeat; Alternative splicing; Polymorphism.
 FT ZN_FING 226 276 PHD-TYPE 1.
 FT ZN_FING 229 274 RING-TYPE.
 FT ZN_FING 273 323 PHD-TYPE 2.
 FT ZN_FING 1102 1155 PHD-TYPE 3.
 FT ZN_FING 1152 1202 PHD-TYPE 4.
 FT ZN_FING 1229 1284 PHD-TYPE 5.
 FT DOMAIN 5121 5242 SET.
 FT DOMAIN 5246 5262 POST-SET.
 FT DOMAIN 2397 2436 COILED COIL (POTENTIAL).
 FT DOMAIN 2788 2809 COILED COIL (POTENTIAL).
 FT DOMAIN 2974 3001 COILED COIL (POTENTIAL).
 FT DOMAIN 3286 3342 COILED COIL (POTENTIAL).
 FT DOMAIN 3437 3476 COILED COIL (POTENTIAL).
 FT DOMAIN 3621 3701 COILED COIL (POTENTIAL).
 FT DOMAIN 4265 4287 COILED COIL (POTENTIAL).
 FT DOMAIN 439 668 15 X 5 AA REPEATS OF S/P-P-P-E/P-E/A.
 FT REPEAT 442 446 1.
 FT REPEAT 460 464 2.
 FT REPEAT 469 473 3.
 FT REPEAT 496 500 4.
 FT REPEAT 504 508 5.
 FT REPEAT 521 525 6.
 FT REPEAT 555 559 7.
 FT REPEAT 564 568 8.
 FT REPEAT 573 577 9.
 FT REPEAT 582 586 10.
 FT REPEAT 609 613 11.
 FT REPEAT 618 622 12.
 FT REPEAT 627 631 13.
 FT REPEAT 645 649 14.
 FT REPEAT 663 667 15.
 FT DOMAIN 229 326 CYS-RICH.
 FT DOMAIN 374 922 PRO-RICH.
 FT DOMAIN 1015 1053 ARG-RICH.
 FT DOMAIN 1122 1235 CYS-RICH.
 FT DOMAIN 1832 2351 PRO-RICH.
 FT DOMAIN 2536 2547 GLN-RICH.

FT	DOMAIN	2587	2703	PRO-RICH.
FT	DOMAIN	2986	4000	GLN-RICH.
FT	DOMAIN	3966	4085	PRO-RICH.
FT	DOMAIN	4634	4702	PRO-RICH.
FT	VARSPPLIC	1	305	Missing (in isoform 2).
FT				/FTId=VSP_008563.
FT	VARSPPLIC	306	672	PMEELPAHSWKCKACRVCRACGAGSAELNPNSEWFENYSLC
FT				HRCHKAQGQGTIRSVAEQHTPVCSRFSPPPEPGDTPPTDEPDA
FT				LYVACQGPQKGGHVTSMQKPEPGPLQCEAKPLGKAGVQLEP
FT				QLEAPLNEEMPLLPPEESPLSPPPEESPTSPPEASRLSP
FT				PPEELPASPLPEALHLSRPLEESPLSPPPEESPLSPPPESS
FT				PFSPLEESPLSPPPEESPPSPALETPLSPPPEASPLSPPFEE
FT				SPLSPPPEELPTSPPEASRLSPPPEESPMSPPEESPMSP
FT				PPEASRLFPPEESPLSPPPEESPLSPPPEASRLSPPPEDS
FT				PMSPPEESPMSPPEEVSRSLPLPVVSRSLSPPPEESPLS
FT				-> MSPPPEESPMSPPEASRLFPPEESPLSPPPEESPLS
FT				PPPEASRLSPPPEDSPMSPPPEESPMSPPEEVSRSLPLPVV
FT				SRLSPPPEESPLSPPPEESPTSPPEASRLSPPPEDSPTSP
FT				PPEDSPASPPPEDSLMSLPLEESPLLPLPEEPQLCPRSEGP
FT				HLSRPEEPHLSRPEEPHLSPOAEHPHLSPOPEEPCLCAV
FT				PEEPHLSPOAEGPHLSPOPEELHLSPOTEEPHLSVPPEPC
FT				LSPQPEESHLSPOSEEPCLSPRPEESHLSPELEKPPLSRPR
FT				EKPPEEPGQCPAPEELPLFPPPGEPSLSPLLGEPALEPGE
FT				PPLSPLPEELPLSPSGEPSLSLQMLPPDPLPPPLSPIITAA
FT				A (in isoform 2).
FT				/FTId=VSP_008559.
FT	VARSPPLIC	1454	1454	E -> EGET (in isoform 3).
FT				/FTId=VSP_008560.
FT	VARIANT	4949	4949	R -> H (in dbSNP:3782356).
FT				/FTId=VAR_017115.
SQ	SEQUENCE	5262 AA;	564171 MW;	26B7C74CAD417E44 CRC64;

Query Match 19.3%; Score 72.5; DB 1; Length 5262;
 Best Local Similarity 34.7%; Pred. No. 1.2e+02;
 Matches 25; Conservative 6; Mismatches 24; Indels 17; Gaps 4;

Qy	12	QVHSSQPSGQAVTVSRQPSLNAVNSLTRSLKRT-----PSLKPDVP----PKP-----	56
		:::: : ::	
Db	2230	ELHAKVPSGQPPNFVRSPGTGAFVG-TPSPMRFTFPQAVGEPKPPVPPQGLPPPHGIN	2288
Qy	57	-SFAPLSTSMKP 67	
Db	2289	SHFGPGPTLGKP 2300	

Search completed: March 24, 2004, 13:14:56
 Job time : 2.43739 secs